AFRL-SR-AR-TR-06-0287

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
05-07-2006	Final	August 2005 - June 2006

4. TITLE AND SUBTITLE

Stochastic Resonance in Signal Detection and Human Perception

5a. CONTRACT NUMBER

FA9550-05-C-0139

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Michels, James H. Chen, Hao
Kay, Steven M. Varshney, Pramod K.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

JHM Technologies, Langmuir Lab, Suite 207, Box 1047, 95 Brown Rd., Ithaca, NY 14850
Syracuse University - 335 Link Hall, Syracuse University, Syracuse, NY 13244

8. PERFORMING ORGANIZATION REPORT NUMBER

JHM - 2005 - 06 - 10

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research
875 North Randolph St., Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Program Manager: Dr. Willard D. Larkin, AFOSR/NL

14. ABSTRACT

Stochastic Resonance (SR) is a nonlinear phenomenon first reported in terms of a nonlinear dynamic effect. The important question of what type of noise to be added has until recently evaded a solution. This issue is addressed directly and a fundamental theoretical framework is developed leading to a determination of the optimal additive SR noise to achieve maximum probability of detection subject to the constraint that the probability of false alarm is not increased. Chapters 2 and 3 provide alternative analytical framework presentations for SR application to detection leading to an optimization solution. Chapter 4 discusses probability of error reduction with implications for communications theory. Subsequent chapters address applications of the analytical theory to suboptimal detectors such as nonparametric detectors (Chapter 5), image enhancement (Chapter 6), and distributed sensor fusion (Chapter 7). Chapter 8 provides a novel consideration to an alternative decision statistic transformation methodology to recover optimal performance for a suboptimal detector. Finally, Chapter 9 presents recommendations and future considerations.

15. SUBJECT TERMS

Stochastic Resonance Noise Enhanced Detection Data Transformation Suboptimal Detector Enhancement Detection Image Enhancement Image Visualization Human Perception CFAR Probability of Error Reduction Parametric Detectors Sensor Fusion Enhancement

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			James H. Michels, JHM Technologies
U	U	U	UU		19b. TELEPHONE NUMBER (Include area code) (607) 257- 5740

STOCHASTIC RESONANCE IN SIGNAL DETECION AND HUMAN PERCEPTION

James H. Michels, Hao Chen, Pramod K. Varshney, and Steven M. Kay

FA9550-050C-0139

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 Introduction

Stochastic Resonance (SR) is a nonlinear phenomenon first reported and analyzed in [Benzi 1981] in terms of a nonlinear dynamic effect. Since then, it was proposed [Benzi 1982] [Nicolis] to explain the observed periodic occurrences of the earth's ice ages using a two state system denoting the earth's climate at present and during the ice age. Variations in the absorption and reflectance of incident solar energy due to changing weather conditions constituted the system 'noise'. The weak periodic 'signal' consisted of variations of incident solar energy due to periodic eccentricity in the earth's orbit. Since then, considerable research efforts have assessed the effect in a wide range of applications including audio systems [Lipshitz], neural networks [Lindner], hyperspectral imaging [Chiang], neuroscience [Kosko], medical imaging [Muller], visual perception [Simonotto], more recently in tactical surveillance [Repperger], as well as applications cited in the reference section.

The classic SR signature is the **signal-to-noise ratio (SNR) gain** of certain nonlinear systems; i.e., the output SNR is significantly higher than the input SNR when an appropriate amount of noise is added. This ratio reflects the gain achieved by the processing procedure. These considerations are treated in references [3] – [17] of [Chen, et. al., 2006b]. Some approaches have been proposed to tune the SR system by maximizing SNR. It has been shown that the SNR of a summing network of excitable units is optimum at a certain level of noise [Collins]. Later, for some SR systems, robustness enhancement using non-Gaussian noise was reported in [Castro, et. al.]. For a fixed type of noise, Mitaim and Kosko [Mitaim, 1998] proposed an adaptive stochastic learning scheme performing a stochastic gradient ascent on the SNR to determine the optimal noise level based on the samples from the process. Rather than adjusting the input noise level, [Xu, et. al.] proposed a numerical method for realizing SR by tuning system parameters to maximize SNR gain. Although SNR is a very important measure of system performance, SR approaches based on SNR gain have several limitations. Specifically, SNR characterizes only the second order terms of the processes; i.e., the signal and noise variance. First, the definition of SNR is not uniform and, in fact, varies from one application to another. Second, to optimize the performance, the complete *a priori* knowledge of the signal is required. Finally, for detection problems where the noise is non-Gaussian, higher order terms may play a role and optimizing output SNR does not guarantee optimizing probability of detection.

SR was also found to enhance the **mutual information (MI)** between input and output signals [Godivier], [Goychuk], [Stocks], [Kosko 2003, 2004], [Mitaim 2004]. Similar to the SNR scenario, for a specified type of SR noise, [Mitaim 2004] showed that almost all noise probability density functions produce some SR effect in threshold neurons and a new statistically robust learning law was proposed to find the optimal noise level. [McDonnell] pointed out that the capacity of a SR channel cannot exceed the actual capacity at the input. Compared to SNR, MI is more directly correlated with the transferred input signal information.

In signal detection theory, SR also plays a very important role in improving the signal detectability. In [Asdi] and [Zozor 2002], improvement of detection performance of a weak sinusoid signal is reported. To detect a DC signal in a Gaussian mixture noise background, [Kay 2000] showed that under certain conditions, performance of the sign detector can be enhanced by

adding some white Gaussian noise. For another suboptimal detector, the locally optimum detector (LOD), [Zozor 2003] pointed out that detection performance is optimum when the noise parameters and detection parameters are matched. A study of the stochastic resonance phenomenon in quantizers conducted in [Saha] showed that improved detection performance can be achieved by a proper choice of the quantizer thresholds. Recently, [Rousseau] pointed out that the detection performance can be further improved by using an optimal detector on the output signal. Despite the progress achieved by the above approaches, the study of the SR effect in signal detection systems is rather limited and does not fully consider the underlying theory. In [Chen 2006b], the underlying mechanism of the stochastic resonance phenomenon was explored for a more general two hypotheses detection problem.

The type of detector that lends itself to improved detection via stochastic resonance is one that nonlinearly processes the data. The eye itself is a nonlinear device and so it is conceivable and has been demonstrated empirically that improved visual detection is possible through this mechanism. The important question of what type of noise to be added has until recently evaded a solution. In Phase I, this issue was addressed directly and a fundamental theoretical concept was developed leading to a determination of the optimal additive SR noise to achieve maximum probability of detection P_D subject to the constraint that the probability of false alarm P_{FA} is not increased.

Clearly, improving visual imagery for human visual perception will depend on the type of nonlinearity that the eye employs. Ultimately, we know that it is the brain that responds to a visual stimulus causing neurons to fire. Conceivably if we understood the effect of the noise PDF on detectability for the "eye detector", then we could add noise such that the total noise PDF is the most desirable one. This requires a study of a.) the types of noise PDFs that can be obtained via convolution since adding noise random variables causes a convolution of their PDFs, b.) the most desirable noise PDF from the standpoint of detectability, and c.) how standard detection theory relates to human visual perception.

We previously pointed out that there are limitations of the signal-to-noise ratio as the most important measure of human detectability. In fact, the SNR measure only characterizes detection in the Gaussian noise case. For non-Gaussian noise, SNR is only part of the story with, for example, the intrinsic accuracy providing the remainder in the independent, identically distributed (IID) non-Gaussian detection problem. The intrinsic accuracy is the single sample Fisher information for a DC level in non-Gaussian IID noise [Kay 1998]. Clearly, these two important aspects of design are related and so an overall strategy of noise PDF design is required. It was also noted in the Phase I proposal that there are also questions of whether the additive enhancement noise should be IID from pixel to pixel. It is possible that correlated noise might be more fruitful. This is an avenue of research that has not been addressed. Furthermore, multivariate PDF design is a difficult problem, but there has been recent progress in this area [Kay 2001], [Tanner 1993], [Ruanaidh 1996], [Gilks 1996]. The above considerations were noted, but not fully considered in the Phase I effort. In Phase II, we shall return to give more thorough consideration to these issues among others.

The establishment of an analytic framework for algorithmic development utilizing stochastic resonance was a prime objective of the Phase I effort. In order to achieve this goal, it was

imperative to assess the relationship of the non-Gaussian nature of the noise processes, the non-linear aspects of the signal processing or suboptimal detection device, as well as the characteristics of the sub-threshold signals.

An important consideration here is that the non-Gaussian noise PDF may actually provide a potential for improved detection performance provided that an appropriate detection strategy is employed. Evidence for this consideration has been noted by [Kay], [Michels] in several publications, although not related to stochastic resonance. Several of these analyses involved partially correlated non-Gaussian noise processes in addition to additive white Gaussian noise. As noted previously, it is possible that the use of correlated noise may be beneficial in the enhancement of the stochastic resonance effect. Both the application of correlated noise as well as the assessment of processes already containing such noise is an open area of research which has received little attention. It remains as an important research topic for the Phase II effort.

In Chapter 2, a novel fundamental theory addressing the optimization of stochastic resonance in detection theory is outlined. The development of this theory was the prime contribution of Phase I. Specifically, this effort achieved the goal of establishing a fundamental analytical framework for the application of stochastic resonance to detection from which an optimization solution could be obtained. Chapter 3 provides an alternative presentation introducing the analytical framework. In Chapter 4, a novel consideration is given to the potential use of an alternative decision statistic transformation methodology to recover optimal performance for a suboptimal detector. Subsequent subsections address the application of the analytical detection theory framework to suboptimal detectors such as nonparametric detectors (Chapter 5), image enhancement (Chapter 6), distributed fusion applications (Chapter 7), and probability of error reduction (Chapter 8) with implications for communications theory. Finally, recommendations and future considerations are presented in Chapter 9.

2.0 A Fundamental Detection Framework using Stochastic Resonance

2.1 Introduction to the Fundamental Detection Framework using SR

This chapter summarizes the mathematical framework developed during Phase I to analyze the stochastic resonance (SR) effect in binary hypothesis testing problems [Chen, et. al., 2006a, b]. Specifically, given an N dimensional data vector $\mathbf{x} \in R^N$, we decide between two hypotheses H_1 and H_0,

$$\begin{aligned} H_0&: \ p_{\mathbf{x}}(\mathbf{x};H_0) = p_0(\mathbf{x}) \\ H_1&: \ p_{\mathbf{x}}(\mathbf{x};H_1) = p_1(\mathbf{x}) \end{aligned} \tag{2.1}$$

where $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$ are the PDFs of $\mathbf{x}$ under H_0 and H_1, respectively. The test above can be completely characterized by a *critical function* (*decision function*) ϕ where $0 \le \phi(\mathbf{x}) \le 1$ for all $\mathbf{x}$. For any observation $\mathbf{x}$, this test chooses the hypothesis H_1 with probability $\phi(\mathbf{x})$. In many cases, $\phi(\mathbf{x})$ can be implicitly expressed by a test statistic $T(\mathbf{x})$ which is a function of $\mathbf{x}$ and a threshold η such that

$$T(\mathbf{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2.2}$$

where its corresponding critical function is

$$\phi_T(\mathbf{x}) = \begin{cases} 1 & T(\mathbf{x}) > \eta \\ \alpha & T(\mathbf{x}) = \eta \\ 0 & T(\mathbf{x}) < \eta \end{cases} \tag{2.3}$$

and $0 \le \alpha \le 1$ is a suitable number. The probability of detection P_D is now given by

$$P_D^{\mathbf{x}} = \int_{R^N} \phi(\mathbf{x}) p_1(\mathbf{x}) d\mathbf{x} \tag{2.4}$$

and the probability of false alarm P_{FA} is given by

$$P_{FA}^{\mathbf{x}} = \int_{R^N} \phi(\mathbf{x}) p_0(\mathbf{x}) d\mathbf{x} \tag{2.5}$$

where the superscripts on P_D and P_{FA} in (2.4) and (2.5) indicate that the test in (2.2) is employed for the data vector $\mathbf{x}$. Although the critical function $\phi(\mathbf{x})$ and test statistic $T(\mathbf{x})$ can take any form, the optimum Neyman-Pearson test involves the likelihood ratio test (LRT) where $T_{LRT}(\mathbf{x}) = p_1(\mathbf{x})/p_0(\mathbf{x})$. Although this test provides optimal P_D subject to the constraint that $P_{FA} = \alpha$ is fixed, the associated LRT requires complete knowledge of the PDFs $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$. In most practical applications, this knowledge is unavailable and may have to be estimated from the data. Also, the input data statistics may very with time or may change from one application to the other. Further, for many detection problems the exact form of the LRT may be too complicated to implement. Therefore, suboptimal detectors featuring simplicity and robustness

are often preferred [Thomas 1970]. To improve a suboptimal detector detection performance, two approaches are widely used. In the first approach, the detector parameters are varied [Zozor 1999, 2002, 2003], [Saha], Galdi]. Alternatively, when the detector itself cannot be altered or the optimum parameters are difficult to obtain, adjusting the observed data becomes a viable approach. Thus, in the work reported here, prime consideration was given to applications in which such detector parameters could not be changed; i.e., the detector parameters were considered to be fixed in terms of both the test statistic and the threshold. This is often the case for applications where the signal processing methodology is not under the user's control. In such cases, we consider the alternative approach of utilizing stochastic resonance.

It is well known that the detection performance can be improved by adding additional noise that is statistically dependent on the existing noise and/or with PDF that depends in which hypothesis is true [Kay 2000]. However, adding a dependent noise is not always possible because pertinent prior information is usually not available. Therefore, we constrain the additive noise to be independent noise. For some suboptimal detectors, as noted in [Kay 2000], detection performance can be improved by adding such noise to the data under certain conditions. For a given type of SR noise, the optimal amount of noise can be determined that maximizes the detection performance for a given suboptimal detector [Inchiosa]. In an effort to explain this noise enhanced phenomenon for some integrate-and-fire neuron models, [Tougaard] demonstrated that the detection performance gain is caused by the nonlinear properties of the spike-generation process itself. However, despite the progress made in the literature, the underlying mechanism of this Stochastic Resonance phenomenon in detection problems has not been fully explored. For example, an interesting problem is the determination of the best 'noise' to be added in order to achieve the best achievable detection performance for the suboptimal detector. In this case, the detection problem can be stated as: Given that the test is fixed, i.e., the critical function $\phi(\cdot)$, as for example T and η, is fixed, can we improve the detection performance by adding SR noise? And if so, what is its PDF to maximize P_D without increasing P_{FA}?

Here, a theoretical analysis is presented to gain further insight into the SR phenomenon and the detection performance of the noise modified observations is obtained. Further, the optimum noise PDF, i.e., not only the noise level but also the noise type is determined. As an illustrative example, the optimum noise PDF as well as several suboptimum noise PDFs are derived for the sign detector. We emphasize that compared to some prior work where one or several nonlinear systems are inserted between the final detector and the original input signal, here, by considering the decision function ϕ in general and the fact that we may consider the entire system between the input signal and output detection resul as a 'super' detector, the results obtained in this paper can be applied for any detection system with any type of fixed SR preprocessing system. Also, compared to the earlier definitions of SR [Benzi 1981], [Gammaitoni], we further extend the concept of 'SR' to a pure noise enhanced phenomenon, ie., a phenomenon of some nonlinear systems in which the system performance is enhanced due to the addition of independent noise at the input. In the work reported here, the terminologies 'SR' and 'noise enhanced' are used interchangeably. However, we point out that the later is actually a generalization of the former.

2.2 Problem Formulation

In order to achieve a possible enhancement of detection performance, we add noise to the original data process **x** and obtain a new data process **y** given by

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{2.6}$$

where **n** is either an independent random process with PDF $p_{\mathbf{n}}(\mathbf{n})$ or a nonrandom signal. Note that here we do not have any constraint on **n**. For example, **n** can be white noise, colored noise, or even a deterministic signal A, corresponding to $p_{\mathbf{n}}(\mathbf{n}) = \delta(\mathbf{n} - A)$. As will be shown later, depending upon the detection problem, an improvement of detection performance may not always be possible. In that case, the optimal noise is equal to zero. The PDF of **y** is expressed by the convolution of the PDFs such that

$$p_{\mathbf{y}}(\mathbf{y}) = p_{\mathbf{x}}(\mathbf{x}) * p_{\mathbf{n}}(\mathbf{x}) = \int p_{\mathbf{x}}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) d\mathbf{x}. \tag{2.7}$$

The binary hypothesis testing problem for this new observed data **y** can be expressed as

$$\begin{cases} H_0 : p_{\mathbf{y}}(\mathbf{y}; H_0) = \int_{R^N} p_0(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) d\mathbf{x} \\ H_1 : p_{\mathbf{y}}(\mathbf{y}; H_1) = \int_{R^N} p_1(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) d\mathbf{x} \end{cases}. \tag{2.8}$$

Since the detector is fixed, i.e., the critical function ϕ of **y** is the same as that for **x**, the probability of detection based on the data **y** is given by,

$$\begin{aligned} P_D^{\mathbf{y}} &= \int_{R^N} \phi(\mathbf{y}) p_y(y; H_1) dy \\ &= \int_{R^N} \phi(\mathbf{y}) \int_{R^N} p_1(\mathbf{x}) p_{\mathbf{n}}(y - x) dx dy \\ &= \int_{R^N} p_1(\mathbf{x}) \left(\int_{R^N} \phi(\mathbf{y}) p_{\mathbf{n}}(y - x) dy \right) dx \\ &= \int_{R^N} p_1(\mathbf{x}) C_{\mathbf{n},\phi}(\mathbf{x}) dx = E_1[C_{\mathbf{n},\phi}(\mathbf{x})] \end{aligned} \tag{2.9}$$

where

$$C_{\mathbf{n},\phi}(\mathbf{x}) \equiv \int_{R^N} \phi(\mathbf{y}) p_{\mathbf{n}}(y - x) dy \tag{2.10}$$

Aternatively,

$$\begin{aligned} P_D^{\mathbf{y}} &= \int_{R^N} p_{\mathbf{n}}(\mathbf{x}) \left(\int_{R^N} \phi(\mathbf{y}) p_1(y - x) dy \right) dx \\ &= \int F_{1,\phi}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{x}) d\mathbf{x} = E_{\mathbf{n}}(F_{1,\phi}(\mathbf{x})) \end{aligned} \tag{2.11}$$

Similarly, we have

$$P_{FA}^{y} = \int_{R^N} p_0(\mathbf{x}) C_{\mathrm{n},\phi}(\mathbf{x}) d\mathbf{x} = E_0[C_{\mathrm{n},\phi}(\mathbf{x})] \tag{2.12}$$

$$= \int F_{0,\phi}(\mathbf{x}) p_{\mathrm{n}}(\mathbf{x}) d\mathbf{x} = E_{\mathrm{n}}(F_{0,\phi}(\mathbf{x})) \tag{2.13}$$

where

$$F_{\mathrm{i},\phi}(\mathbf{x}) \equiv \int_{R^N} \phi(\mathbf{y}) p_i(\mathbf{y}-\mathbf{x}) d\mathbf{y} \quad \mathrm{i} = 0,\ 1, \tag{2.14}$$

and $F_{\mathrm{i},\phi}(\mathbf{x})$ corresponds to hypothesis H_{i}, and $E_{\mathrm{i}}(\cdot)$, $E_{\mathrm{n}}(\cdot)$ are the expected values based on the distributions $\mathrm{p_i}(\cdot)$ and $\mathrm{p_n}(\cdot)$, respectively. Note that $F_{\mathrm{i}}(\mathbf{x})$ has the property that $\mathrm{P_{FA}^x} = F_{0,\phi}(0)$ and $\mathrm{P_D^x} = F_{1,\phi}(0)$. To simplify notation, we omit the subscript ϕ of F and C and denote them as F_1, F_0 and C_{n}, respectively. Further, from (2.14), $F_1(\mathbf{x}_0)$ and $F_0(\mathbf{x}_0)$ are actually the probability of detection and probability of false alarm, respectively, for this detection scheme with input $\mathbf{y} = \mathbf{x} + \mathbf{x}_0$. For example, $F_1(-2)$ is the P_D of this detection scheme with input $\mathbf{x} - 2$. Therefore, it is very convenient for us to obtain the F_1 and F_0 values by analytic computation of values by analytic computation if p_0, p_1 and ϕ are known. When they are not available, F_1 and F_0 can be obtained from the data itself by processing it through the detector and recording the detection performance. Thus, it is not necessary to have complete knowledge regarding $\phi(\cdot)$ and $p_{\mathrm{i}}(\cdot)$. From (2.11) and (2.13), we may formalize the optimal SR noise definition as follows.

Consider the two hypothesis detection problem as in (2.1). The PDF of the optimum SR noise is given by

$$\mathrm{p_n^{opt}} = \arg\max_{\mathrm{p_n}} \int_{R^N} F_1(\mathbf{x}) \mathrm{p_n}(\mathbf{x}) d\mathbf{x} \tag{2.15}$$

where

1) $\mathrm{p_n}(\mathbf{x}) \geq 0, \mathbf{x} \in R^N$

2) $\int_{R^N} \mathrm{p_n}(\mathbf{x}) d\mathbf{x} = 1$

3) $\int_{R^N} F_0(\mathbf{x}) \mathrm{p_n}(\mathbf{x}) d\mathbf{x} \leq F_0(0)$.

Conditions 1) and 2) are fundamental properties of a PDF function. Condition 3) ensures that $P_{FA}^{y} \leq P_{FA}^{x}$, i.e., the P_{FA} constraint under the Neyman-Pearson criterion is satisfied. Further, if the inequality condition in 3) becomes equality, the constant false alarm rate (CFAR) property of the original detector is maintained.

2.3 Optimum SR Noise for Neyman-Pearson Detection

In general, it is difficult to find an exact form of $p_{\mathrm{n}}(\cdot)$ directly because of condition 3). However, an alternative approach considers the relationship between $\mathrm{p_n}(\mathbf{x})$ and $F_{\mathrm{i}}(\mathbf{x})$. From (2.14), for a given value f_0 of F_0, we have $\mathbf{x} = F_0^{-1}(f_0)$, where F_0^{-1} is the inverse of F_0. When F_0 is a one-to-one mapping function, $\mathbf{x}$ is a unique vector. Otherwise, $F_0^{-1}(f_0)$ is a set of $\mathbf{x}$ for which $F_0(\mathbf{x}) = f_0$. Therefore, we can express a value or a set of values f_1 of F_1 as

$$f_1 = F_1(\mathbf{x}) = F_1(F_0^{-1}(f_0)) . \tag{2.16}$$

Given the noise distribution of $p_{\mathbf{n}}(\cdot)$ in the original R^N domain, $p_{\mathbf{n},f_0}(\cdot)$, the noise distributiom in the f_0 domain can also be uniquely determined. Further. the conditions on the optimum noise can be rewritten in terms of f_0 equivalently as

4) $\mathrm{p}_{\mathbf{n},f_0}(f_0) \geq 0$

5) $\int \mathrm{p}_{\mathbf{n},f_0}(f_0) df_0 = 1$

6) $\int f_0 \mathrm{p}_{\mathbf{n},f_0}(f_0) df_0 \leq \mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}$

and

$$\mathrm{P}_{\mathrm{D}}^{\mathrm{y}} = \int_0^1 f_1 p_{\mathbf{n},f_0}(f_0) df_0 , \tag{2.17}$$

where $\mathrm{p}_{\mathbf{n},f_0}(f_0)$ is the SR noise PDF in the f_0 domain. Compared to the original conditions 1), 2) and 3), this equivalent form has some advantages. First, the problem complexity is dramatically reduced. Rather than searching for an optimal solution in R^N, an optimal solution is sought in a single dimensional space. Second, by applying these new conditions, we avoid the direct use of the underlying PDFs $\mathrm{p}_1(\cdot)$ and $\mathrm{p}_0(\cdot)$ and replace them with f_1 and f_0, respectively. Note that in some cases, it is not easy to find the exact form of f_1 and f_0. However, recall that $F_1(\mathbf{x}_0)$ and $F_0(\mathbf{x}_0)$ are the probability of detection and false alarm, respectively, of the original system $\mathbf{x} + \mathbf{x}_0$. In practical applications, we may learn the relationship by Monte Carlo simulation using importance sampling. In general compared to p_1 and p_0, f_1 and f_0 are much easier to estimate and once the optimum $\mathrm{p}_{\mathbf{n},f_0}$ is found, the optimum $p_{\mathbf{n}}(\mathbf{x})$ is determined as well by virtue of the inverse of the function F_0 and F_1.

Let us now consider the function $J(\mathrm{t})$ such that $J(\mathrm{t}) = \sup(f_1 : f_0 = \mathrm{t})$ is the maximum value of f_1 given f_0. Clearly, $J(P_{FA}^x) \geq F_1(0) = P_D^x$. From (2.17), it follows that for any noise $p_{\mathbf{n}}$, we have

$$\mathrm{P}_{\mathrm{D}}^{\mathrm{y}}(p_{\mathbf{n}}) = \int_0^1 J(f_0) p_{\mathbf{n},f_0}(f_0) df_0 . \tag{2.18}$$

Therefore, the optimum P_D^y is attained when $f_1(f_0) = J(f_0)$ and $P_{D,opt}^y = E_{\mathbf{n}}(J)$.

A. Determination of SR Detection Improvement

A significant contribution of the theoretical framework provided here is the capability to determine if SR will indeed provide performance improvement in a given problem. Detection enhancement results using SR results if $P_{D,opt}^y > P_D^x$. However, it requires complete knowledge of $F_0(\cdot)$ and $F_1(\cdot)$ as well as significant computation. For a large class of detectors, however,

depending on the specific properties of J, we may determine the sufficient conditions for improvability and non-improvability more readily. These are given in the following theorems the proofs of which are contained in [Chen 2006b].

Theorem 1 (Improvability of Detection via SR): If $J(\mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}) > \mathrm{P}_D^x$ or $J''(\mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}) > 0$ when $J(\mathrm{t})$ is second order continuously differentiable around $\mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}$, then there exists at least one noise process $\mathbf{n}$ with PDF $p_{\mathbf{n}}(\cdot)$ that can improve the detection performance.

Theorem 2 (Non-improvability of Detection via SR): If there exists a non-decreasing concave function $\psi(f_0)$ where $\psi(\mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}) = J(\mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}) = F_1(0)$ and $\psi(f_0) \geq J(f_0)$ for every f_0, then $\mathrm{P}_D^y \leq \mathrm{P}_D^x$ for any independent noise, i.e., the detection performance cannot be improved by adding noise.

B. Determination of the Optimum SR Noise PDF

Another very significant result of the theoretical framework is the determination of the exact form of $\mathrm{p}_{\mathrm{n}}^{opt}$ subject to the constraint $\mathrm{P}_{\mathrm{FA}}^{\mathrm{y}} \leq \mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}$. This is contained in Theorem 3 [Chen 2006b].

Theorem 3 (Form of Optimum SR Noise): To maximize P_D^y under the constraint $\mathrm{P}_{\mathrm{FA}}^{\mathrm{y}} \leq \mathrm{P}_{\mathrm{FA}}^{\mathrm{x}}$, the optimum noise can be expressed as[1]

$$\mathrm{p}_{\mathrm{n}}^{opt}(\mathbf{n}) = \lambda\delta(\mathbf{n}-\mathbf{n}_1) + (1-\lambda)\delta(\mathbf{n}-\mathbf{n}_2) \tag{2.19}$$

where $0 \leq \lambda \leq 1$. Specifically, to obtain the maximum achievable performance given the false alarm constraint, the optimum noise is a randomization of two discrete vectors added with probability λ and $(1 - \lambda)$, respectively. It can also be shown [Chen 2006b] that

$$\mathrm{p}_{\mathbf{n},f_0}^{opt} = \lambda\delta(f_0 - f_{01}) + (1-\lambda)\delta(f_0 - f_{02}) \tag{2.20}$$

where $f_{01} = F_0(\mathbf{n}_1)$ and $f_{02} = F_0(\mathbf{n}_2)$. Alternatively, the optimum SR noise can also be expressed in terms of $C_{\mathbf{n}}$, such that

$$C_{\mathbf{n}}^{opt}(x) = \lambda\phi(\boldsymbol{x}+\mathbf{n}_1) + (1-\lambda)\phi(\boldsymbol{x}+\mathbf{n}_2)\,. \tag{2.21}$$

From (2.20), we have

$$\mathrm{P}_{D,opt}^{y} = \lambda J(f_{01}) + (1-\lambda)J(f_{02}) \tag{2.22}$$

and

$$\mathrm{P}_{FA,opt}^{y} = \lambda f_{01} + (1-\lambda)f_{02} \leq \mathrm{P}_{FA}^{x}\,. \tag{2.23}$$

[1] This form of optimum noise PDF is not necessarily unique. There may exist other forms of noise PDF that achieve the same detection performance.

C. Determination of the PDF of Optimum SR Noise

Depending upon the location of the maxima of $J(\cdot)$, we have the following theorem.

Theorem 4 Let $F_{1M} = \max(J(t)) = t_0$ and $t_0 = \arg\min_t(J(t) = F_{1M})$. It follows that

Case 1: If $t_0 \le P_{FA}^x$, then $P_{FA,opt}^y = t_0$ and $P_{D,opt}^y = F_{1M}$, i.e., the maximum achievable performance is obtained when the optimum noise is a DC signal with value $\mathbf{n}_0$ or

$$p_n^{opt}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_0) \tag{2.24}$$

where $F_0(\mathbf{n}_0) = t_0$ and $F_1(\mathbf{n}_0) = F_{1M}$. Here, the maximum probability of detection is achieved subject to the probability of false alarm constraint $P_{FA}^y \le P_{FA}^x$ by adding a constant to the input with a value that depends upon the decision regions and the probability density functions under the two hypotheses. However, the threshold must be varied to maintain $P_{FA}^y \le P_{FA}^x$.

Case 2: If $t_0 > P_{FA}^x$, then $P_{FA,opt}^y = F_0(0) = P_{FA}^x$, i.e., the inequality of (2.23) becomes equality. Furthermore,

$$P_{FA,opt}^y = \lambda f_{01} + (1-\lambda) f_{02} = P_{FA}^x . \tag{2.25}$$

In this case, the probability of false alarm is maintained without a threshold variation. Thus, the CFAR property is achieved.

2.4 A Detection Problem Example

The above theoretical methodology was applied to the problem considered by [Kay 2000]. Given observation data **x**[n], n = 1, 2, …, N, we consider the binary hypothesis testing problem such that

$$\begin{cases} H_0 : \mathrm{x[n]} = \mathrm{w[n]} & \mathrm{n} = 1, 2, ..., \mathrm{N} \\ H_1 : \mathrm{x[n]} = A + \mathrm{w[n]} & \mathrm{n} = 1, 2, ..., \mathrm{N} \end{cases} \tag{2-26}$$

involving detection of a known dc signal level A > 0 in i.i.d. symmetric Gaussian mixture noise w[n] with PDF

$$p_w(\mathrm{w}) = \frac{1}{2}\gamma(w; -\mu, \sigma_0^2) + \frac{1}{2}\gamma(w; \mu, \sigma_0^2) \tag{2.27}$$

where

$$\gamma(w; \mu, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(w-\mu)^2}{2\sigma^2} \right]. \tag{2.28}$$

Here, $\mu = 3$, $A = 1$ and $\sigma_0 = 1$. The suboptimal sign detector is considered with test statistic

$$T(\mathbf{x}) = \frac{1}{N}\sum_{i=0}^{N-1}\left(\frac{1}{2}+\frac{1}{2}\mathrm{sgn}(x[i]\right) = \frac{1}{N}\sum_{i=0}^{N-1}\left(\varpi_x[i]\right). \tag{2.29}$$

where $\varpi_x[i] = \frac{1}{2}+\frac{1}{2}\mathrm{sgn}(x[i])$ From the second equality in (2.29), we consider this detector as essentially the fusion of the decision results from N i.i.d. sign detectors. For N = 1, the detection problem reduces to a problem with test statistic $T_1(x) = x$, threshold $\eta = 0$ (sign detector), and probability of false alarm $P_{FA}^x = 0.5$. The distribution of x under the H_0 and H_1 hypotheses can be expressed as

$$p_0(x) = \frac{1}{2}\gamma(x;-\mu,\sigma_0^2)+\frac{1}{2}\gamma(x;\mu,\sigma_0^2) \tag{2.30}$$

and

$$p_1(x) = \frac{1}{2}\gamma(x;-\mu+A,\sigma_0^2)+\frac{1}{2}\gamma(x;\mu+A,\sigma_0^2), \tag{2.40}$$

respectively. The critical function is given by

$$\phi(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases}. \tag{2.41}$$

The problem of determining the optimal SR noise is to find the optimal $p_n(n)$ where for the new observation $y = x + n$, the probability of detection $P_D^y = p(y > 0; H_1)$ is maximum while for the probability of false alarm, $P_{FA}^y = p(y > 0; H_0) \le P_{FA}^x = \frac{1}{2}$. It follows [Chen 2006b] that the resulting optimal SR noise PDF is

$$p_n^{opt}(n) = \lambda\delta(n - n_1) + (1-\lambda)\delta(n - n_2) = 0.3085\delta(n + 3.5) + 0.6915\delta(n - 2.5). \tag{2.42}$$

It also follows that for the case of optimal SR noise, but now constrained to have a symmetric PDF, $p_s^{opt}(x)$, where $p_s^{opt}(x) = p_s^{opt}(-x)$, we have for the example considered here

$$p_s^{opt}(x) = \frac{1}{2}\delta(x-\mu)+\frac{1}{2}\delta(x+\mu). \tag{2.43}$$

Performance results are shown in Fig. 2.1. First, let us consider Fig. 2.1a which plots the curve $U = (f_1, f_0)$ where $f_1 = F_1(\mathbf{x})$ and $f_0 = F_0(\mathbf{x})$. This curve is significant in that it reveals the potential for detection performance improvement via SR. This is observed by again noting that for the original data process $\mathbf{x}$, $P_{FA}^x = F_0(0)$ and $P_D^x = F_1(0)$. For the problem considered here, $P_{FA}^x = F_0(0) = f_0 = 0.5$ and $P_D^x = F_1(0) = f_1 = 0.51$ yielding detection probability barely above P_{FA}. However, as indicated by the curve that upper bounds the convex hull, V, which contains

all possible P_D and P_{FA} values after SR is added, $P_D^y = 0.6915$. Thus, the region for potential performance improvement via SR is the region of the convex hull V that lies above the curve U = (f_1, f_0). Further, this indicates that if the curve U is a convex function, there is no potential for performance improvement via SR. Thus, the curve U = (f_1, f_0) provides an important diagnostic tool for the determination of potential performance enhancement.

Fig. 2.1b plots P_D^y versus the signal level A. The lowest curve plots P_D^x with no SR noise. As expected, it increases with increasing signal level. The next three curves, which lie just above, utilized the optimum PDFs for three symmetric SR noise cases. Ranging from lowest to highest, the SR noise consisted of white Gaussian, uniform, and the optimal symmetric SR noise of (2.43), respectively. We observe that for the symmetric SR noise cases, the curves all converge at A = μ to a common P_D^y value which lies on the P_D^x curve. Thus, for values of A > μ, no performance improvement is obtained using symmetric SR noise. However, for small values of A, say 0 < A < 0.6, the P_D^y values of the symmetric optimal SR noise case achieves the same level of performance as those obtained for the case of optimal SR noise with PDF given by (2.42). For A > 0.6, however, the latter optimal noise case achieves superior performance levels closer to the optimal likelihood ratio test (LRT).

Fig. 2.1c considers the detection performance dependence upon the background noise standard deviation, σ_0. The lowest curve plots P_D^x versus σ_0; i.e., the case for which no SR noise is added. We observe that as σ_0 increases, P_D^x increases until the noise standard deviation reaches the level $\sigma_1 = 2.942$. For low values of σ_0 (high SNR), however, the optimum SR noise enhanced detector reaches $P_D^y \approx 1$, while for the symmetric SR noise enhanced detectors, the performance is reduced. As σ_0 increases, the performance of the SR enhanced detectors converges to the P_D^x value at $\sigma_0 = \sigma_1$. This results from the convergence of the bimodal Gaussian background noise PDF $p_0(\mathbf{x})$ to a unimodal PDF as σ_0 approaches σ_1. At this point, the decision function $\phi(\mathbf{x})$ and the LRT test are equivalent for $P_{FA} = 0.5$. Thus, adding SR noise will not improve P_D and all the detection results converge to P_D^x.

Fig. 2.1d shows each detector's performance with respect to μ for A = 1 and $\sigma_0 = 1$. All the detectors show a decrease in performance as μ increases from zero. However, for $\mu > \mu_0 \approx 1.5$, the optimum LRT begins to increase with increasing μ. To explain this effect, we note that for $\mu << \mu_0$, the bimodal Gaussian noise is approximately unimodal. However, as these two Gaussian mixture peaks separate, the detectability initially decreases until the peaks are sufficiently separated. As $\mu \rightarrow \infty$, the peak separation is sufficient such that the background noise PDF is essentially unimodal for signal level A > 0.

Finally, Fig. 2.2 shows the ROC curve for this problem, but now with N = 30. Again, the superior performance of the optimal SR noise enhanced detectors is observed compared with the cases of uniform and Gaussian SR noise. Specifically, the detection performance is much closer to the optimal LRT curve.

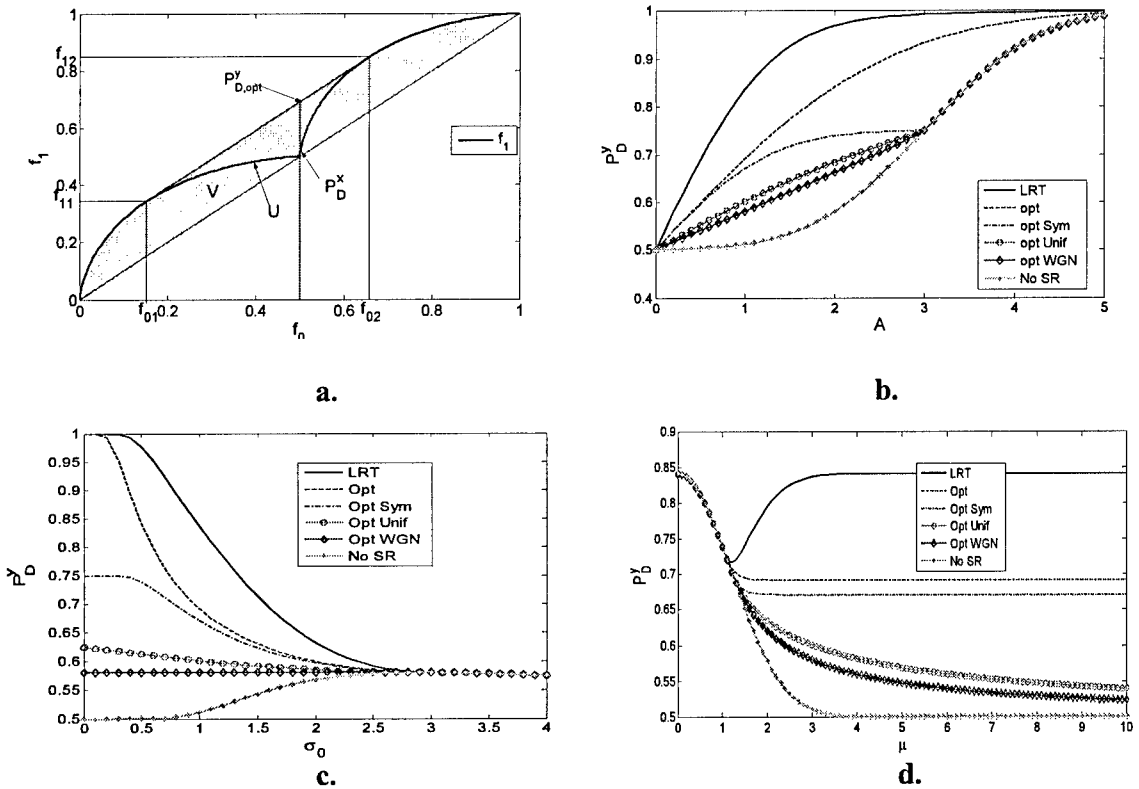


Fig. 2.1 a.) Plot of U = (f_1, f_0) where $f_1 = F_1(\mathbf{x})$ and $f_0 = F_0(\mathbf{x})$, b.) P_D^y versus signal level A, c.) P_D^y versus σ_0, d.) P_D^y versus μ.

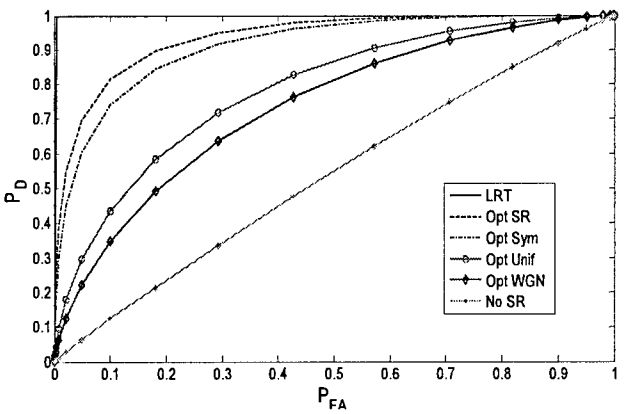


Fig. 2.2 Receiver operating characteristic (ROC), P_D^y versus P_{FA}^y .

3.0 An Alternative Derivation for the SR Detection Framework

The following derivation [Kay, 2005] addresses the problem of deciding between two hypotheses based on a single sample. The sample may be a single data sample or a test statistic. The scalar test statistic is x, which under H_0 has PDF $p_0(x)$ and under H_1, $p_1(x)$. It is assumed that we decide H_1 if $x > 0$. Now consider the same detector but replace x by $y = x + c$, where c is a random variable *that is independent of* x. Then, we decide H_1 if $y = x + c > 0$. The PDF of c is $p_C(c)$ and it is this PDF that is of interest. We allow the use of impulses in the PDF so that c may be either continuous, discrete or mixed variable. For this detector, the probability of false alarm and probability of detection are

$$P_{FA} = \int_0^\infty p_0^Y(y)dy \tag{3.1}$$

$$P_D = \int_0^\infty p_1^Y(y)dy \tag{3.2}$$

where the PDFs of y are

$$p_0^Y(y) = \int_{-\infty}^{\infty} p_0(y-c)p_C(c)dc$$

$$p_1^Y(y) = \int_{-\infty}^{\infty} p_1(y-c)p_C(c)dc$$

which follows from the usual result for the sum of two independent random variables. Explicitly then

$$\begin{aligned} P_{FA} &= \int_0^\infty \left[\int_{-\infty}^{\infty} p_0(y-c)p_C(c)dc \right] dy \\ &= \int_{-\infty}^{\infty} \int_0^\infty p_0(y-c)dy p_C(c)dc \\ &= \int_{-\infty}^{\infty} P_{FA}(c)p_C(c)dc \end{aligned}$$

where $P_{FA}(c) = \int_0^\infty p_0(y-c)dy$ denotes the *conditional probability of false alarm.* It is the probability of false alarm conditional on observing C = c. Note that the unconditional probability of false alarm P_{FA} is just $E_C[P_{FA}(C)]$, where E_C denotes expectation with respect to the PDF $p_C(c)$. Similarly, we have

$$P_D = \int_{-\infty}^{\infty} P_D(c)p_C(c)dc = E_C[P_D(C)]$$

where

$$P_D(c) = \int_0^\infty p_1(y-c)dy .$$

It is interesting to note that $P_D(c)$ is just the probability of detection for the original detector but with the threshold, originally given by zero, replaced by –c. This is because we decide H_1 if $y = x + c > 0$ or equivalently if $x > -c$. The threshold, however, is a random variable and hence the overall detector performance is given by the expected value of $P_D(c)$.

Our problem has now reduced to the following. Choose $p_C(c)$ so that $E_C[P_D(C)]$ is maximized subject to the constraint that $E_C[P_{FA}(C)] = ½$. (We assume continuity of the original ROC so that the false alarm constraint is an equality). To proceed further it is useful to simplify $P_{FA}(c)$ and $P_D(c)$. Consider

$$\begin{aligned} P_{FA}(c) &= \int_0^\infty p_0(y-c)dy \\ &= \int_{-c}^\infty p_0(t)dt \quad (\text{let } t = y - c). \end{aligned}$$

Now note that in this form it is clear that $P_{FA}(c)$ is just the complementary distribution function or right-tail-probability. Also, as such, it is obvious that as c increases (-c decreases), that $P_{FA}(c)$ also increases. Thus, $P_{FA}(c)$ is monotonically increasing with c. It is well known that if a function is monotonically increasing, then the inverse function exists and it too is monotonically increasing; for example, $g(x) = \exp(x)$. Thus, we will use a variable change by letting $u = P_{FA}(c)$, where if $-\infty < c < \infty$, we must have $0 \le u \le 1$. Also the inverse function is denoted as $c = P_{FA}^{-1}(u)$. Hence, we now have the new random variable $U = P_{FA}(C)$ and therefore

$$P_{FA} = E_C[P_{FA}(C)] = E_U[U] = \int_0^1 u p_U(u)du$$

and similarly

$$P_D = E_C[P_D(C)] = E_U[P_D(P_{FA}^{-1}(U))] = \int_0^1 P_D(P_{FA}^{-1}(u))p_U(u)du .$$

Recall that we desire $P_{FA} = 1/2$. Thus, the equivalent optimization problem is to maximize

$$J(p_U) = \int_0^1 P_D(P_{FA}^{-1}(u))p_U(u)du \qquad (3.3)$$

subject to the constraint that

$$\int_0^1 u p_U(u)du = ½.$$

However, it is more convenient to write the constraint as

$$\int_0^1 (u-1/2) p_U(u)du = 0 = E_U[U - ½]$$

and to define a new random variable W = U – ½ so that the constraint becomes $E_W[W] = 0$. Then, $J(p_U)$ becomes from (3.3)

$$J(p_U) = E_U[P_D(P_{FA}^{-1}(U))] = E_W[P_D(P_{FA}^{-1}(W+1/2))] = J(p_W)$$

where explicitly

$$J(p_W) = \int_{-1/2}^{1/2} P_D(P_{FA}^{-1}(w+1/2)) p_W(w)dw$$
$$= \int_{-1/2}^{1/2} g(w) p_W(w)dw$$

and

$$g(w) = P_D(P_{FA}^{-1}(w+1/2)).$$

Summarizing, we wish to maximize over $p_W(w)$ the functional

$$\int_{-1/2}^{1/2} g(w) p_W(w)dw$$

subject to the constraint that $E_W[W] = 0$. Note that the random variable $W = P_{FA}(C) – ½$ takes on values in the interval [-1/2, 1/2]. We can further simplify the problem by maximizing the functional

$$\int_{-1/2}^{1/2} (g(w) - (w+1/2)) p_W(w)dw$$

since $\int_{-1/2}^{1/2} (w+1/2)) p_W(w)dw = ½$ due to the $E_W[W] = 0$ constraint. Letting

$$h(w) = g(w) - (w + ½)$$

which is explicitly

$$h(w) = P_D(P_{FA}^{-1}(w+1/2)) - (w + ½) \qquad (3.4)$$

we finally seek to maximize

$$\int_{-1/2}^{1/2} h(w) p_W(w) dw$$

subject to the constraint that $E_W[W] = 0$. Note that the function h(w) takes on nonnegative values if $P_D(c) \geq P_{FA}(c)$ (the ROC for a variable threshold of the detector that decides H_1 if $x > -c$ is above the 45^0 line). This is because

$$h(w) = P_D(P_{FA}^{-1}(w+1/2)) - (w + \tfrac{1}{2}) = P_D(P_{FA}^{-1}(u)) - u = P_D(c) - P_{FA}(c) \geq 0.$$

Also, at the end points of the [-1/2, ½] interval in w we have

$$h(-1/2) = P_D(P_{FA}^{-1}(0)) = P_D(+\infty) = 0$$
$$h(1/2) = P_D(P_{FA}^{-1}(1)) - 1 = P_D(-\infty) - 1 = 0.$$

A typical plot of $h(w)$ is shown in Fig. 3.1. This example will be used later.

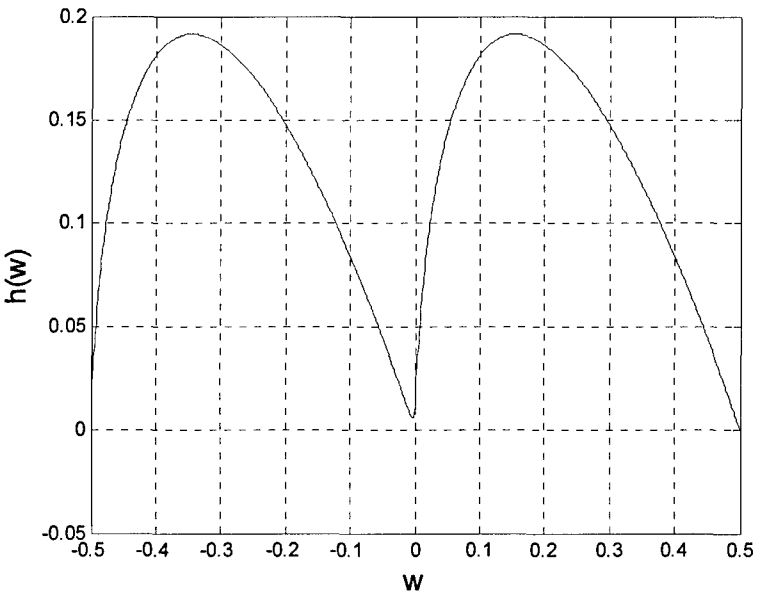


Fig. 3.1 Plot of the $h(w)$ function for the example problem.

In essence, we have modified the function F_1 versus F_0 in Section 2.0 so that it is plotted as the difference between the function and the 45^0 line. (See also Fig. 2.1a for F_1 versus F_0). Also, it has been shifted to be defined over the symmetric interval [-1/2, 1/2]. This has the advantage of simplifying extensions of the results given here. In addition, the maxima, which will be required later, are easily found numerically using efficient routines such as a golden search.

The problem now is to maximize the functional

$$J(p_W) = \int_{-1/2}^{1/2} h(w) p_W(w) dw \tag{3.5}$$

subject to the constraint that $E_W[W] = 0$. The function $h(w)$ is nonnegative over the interval [-1/2, 1/2]. To simplify the discussion we will assume that $h(w)$ has a unique maximum over the interval (0,1/2), and a unique maximum over the interval [-1/2, 0]. The maximum is assumed not to occur at $w = 0$. Also, it is assumed that the maximum values are equal. These assumptions are satisfied for the example given in [Kay, 2000]. (More rigorous and general results can be obtained using standard theorems in analysis such as 'continuous functions on compact sets', etc.). Once the optimal p_W is found, the PDF for C or p_C can be found by transforming back to C using the relationship $W = P_{FA}(C) - \frac{1}{2}$ and the standard results in transformation of random variables. Note that if the maximum of $h(w)$ over the interval [-1/2, 1/2] were to occur at $w = 0$, then the constraint $E_W[W] = 0$ would be satisfied for $p_W(w) = \delta(w)$ and $J(p_W)$ would also be maximized. Thus, the solution would be to choose $c = P_{FA}^{-1}(w+1/2) = P_{FA}^{-1}(1/2) = 0$ and no improvement in performance would be possible.

Continuing, we let w_- be the value that maximizes $h(w)$ for $w < 0$ and w_+ be the value that maximizes $h(w)$ for $w > 0$ ($w = 0$ is excluded), and assume that $h(w_-) = h(w_+)$. Let the set A denote the remaining portion of the interval [-1/2, 1/2] so that $\{ w_-, w_+\} \cup A = [-1/2, 1/2]$ and $\{w_-, w_+\} \cap A = \varnothing$. Next, represent $p_W(w)$ as

$$p_W(w) = \alpha(w_-)\delta(w - w_-) + \alpha(w_+)\delta(w - w_+) + p_W(w) I_A(w) \tag{3.6}$$

where δ denotes the Dirac delta function and $I_A(w) = 1$ for $w \in A$ and is zero otherwise (the indicator function). Actually, any PDF may be decomposed this way subject to the condition that

$$\alpha(w_-) + \alpha(w_+) + \int_A p_W(w) dw = 1.$$

Using this in (3.5) produces

$$J(p_W) = h(w_-)\alpha(w_-) + h(w_+)\alpha(w_+) + \int_A h(w) p_W(w) dw. \tag{3.7}$$

In order to satisfy the constraint, we must have that $p_W(w)$ has mass for $w > 0$ and $w < 0$ (or the probability of W being negative is nonzero and the probability of W being positive is also nonzero). Otherwise, we could not have $E_W[W] = 0$. The constraints on $p_W(w)$ form the two linear equations

$$\alpha(w_-) + \alpha(w_+) + \int_A p_W(w) dw = 1$$
$$\alpha(w_-) w_- + \alpha(w_+) w_+ + \int_A w p_W(w) dw = 0.$$

In matrix form the constraints are

$$\mathbf{A}\alpha = \mathbf{b}$$

or

$$\begin{bmatrix} 1 & 1 \\ w_- & w_+ \end{bmatrix} \begin{bmatrix} \alpha(w_-) \\ \alpha(w_+) \end{bmatrix} = \begin{bmatrix} 1 - \int_A p_W(w)dw \\ -\int_A w p_W(w)dw \end{bmatrix} \tag{3.8}$$

where **A**, α and **b** are defined implicitly. With $\mathbf{h} = [h(w_-)\ h(w_+)]^T$, we have from (3.7)

$$\begin{aligned} J(p_W) &= \mathbf{h}^T\alpha + \int_A h(w)p_W(w)dw \\ &= \mathbf{h}^T\mathbf{A}^{-1}\mathbf{b} + \int_A h(w)p_W(w)dw \\ &= [h(w_-)\ h(w_+)]\frac{\begin{bmatrix} w_+ & -1 \\ -w_- & 1 \end{bmatrix}}{w_+ - w_-}\begin{bmatrix} 1 - \int_A p_W(w)dw \\ -\int_A w p_W(w)dw \end{bmatrix} + \int_A h(w)p_W(w)dw \\ &= \begin{bmatrix} \frac{w_+h(w_-) - w_-h(w_+)}{w_+ - w_-} & \frac{h(w_+) - h(w_-)}{w_+ - w_-} \end{bmatrix}\begin{bmatrix} 1 - \int_A p_W(w)dw \\ -\int_A w p_W(w)dw \end{bmatrix} + \int_A h(w)p_W(w)dw. \end{aligned}$$

We point out that the two terms in the first bracket can be expressed as

$$\frac{w_+h(w_-) - w_-h(w_+)}{w_+ - w_-} = h(w_+)$$

and

$$\frac{h(w_+) - h(w_-)}{w_+ - w_-} = 0.$$

Recalling that $h(w_-) = h(w_+)$ and recognizing that $h(w) < h(w_+)$ for all $w \in A$, we have that

$$\begin{aligned} J(p_W) &= h(w_+) + \int_A (h(w) - h(w_+))p_W(w)dw \\ &\le h(w_+). \end{aligned}$$

Clearly, the upper bound is attained when $p_W(w) = 0$ for all $w \in A$. This results in the solution from (3.6) of

$$p_W(w) = \frac{w_+}{w_+ - w_-}\delta(w - w_-) + \frac{w_-}{w_- - w_+}\delta(w - w_+) \tag{3.9}$$

where we have solved for $\alpha(w_-)$ and $\alpha(w_+)$ by using (3.8) with the right-hand-side vector being $[1\ 0]^T$. Since W can only take on values w_- and w_+, it follows that the only values C can take on are

$$c_- = P_{FA}^{-1}(w_- + 1/2)$$
$$c_+ = P_{FA}^{-1}(w_+ + 1/2). \tag{3.10}$$

The optimal PDF for C is therefore

$$p_C(c) = \frac{w_+}{w_+ - w_-}\delta(c - c_-) + \frac{w_-}{w_- - w_+}\delta(c - c_+) \tag{3.11}$$

where c_- and c_+ are given by (3.10) and w_- and w_+ are the values that maximize the $P_D(P_{FA}^{-1}(w+1/2)) - (w+1/2)$ for $-1/2 \le w < 0$ and $0 < w \le ½$, respectively.

We again consider the example [Kay 2000] of Section 2.0 and assume only one sample as described above. The PDFs are given by

$$p_0(x) = \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x-3)^2\right) + \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x+3)^2\right)$$
$$p_1(x) = p_0(x - 1).$$

Since for a fixed c we decide H_1 if $x > -c$, we have

$$P_{FA}(c) = \frac{1}{2}Q(-c-3) + \frac{1}{2}Q(-c+3)$$
$$P_D(c) = \frac{1}{2}Q(-c-4) + \frac{1}{2}Q(-c+2)$$

where $Q(x)$ is the right-tail probability for a standard normal random variable. This is plotted in Fig. 3.2. The function $h(w)$ is given by (3.4) as

$$h(w) = P_D(P_{FA}^{-1}(w+1/2)) - (w+1/2)$$

and can be evaluated over [-1/2, 1/2] as in Fig. 3.1. A numerical search finds the maxima of $h(w)$, which when converted to c yields $c_- = -3.507$ and $c_+ = 2.493$. The other values become

$$\frac{w_+}{w_+ - w_-} = 0.694 \quad \text{and} \quad \frac{w_-}{w_- - w_+} = 0.306$$

so that the optimal PDF is

$$p_C(c) = 0.306\delta(c + 3.507) + 0.694\delta(c - 2.493)$$

in agreement with the results obtained in Section 2.0.

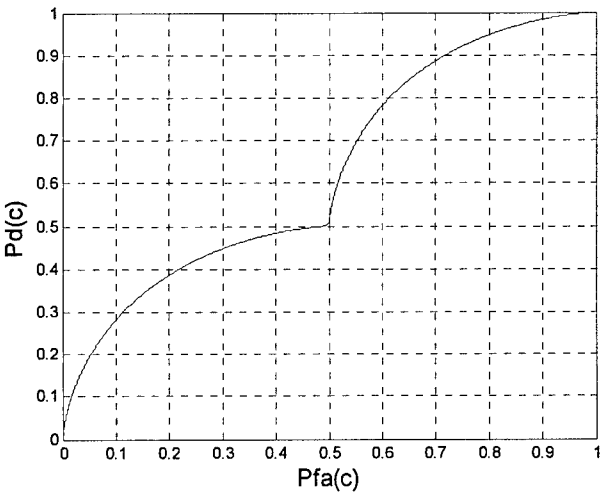


Fig. 3.2 Plot of $P_D(c)$ versus $P_{FA}(c)$.

4.0 Reducing Probability of Decision Error using Stochastic Resonance

In this chapter, we address the problem of reducing the probability of decision error of an existing binary receiver that is suboptimal using the ideas of stochastic resonance. The optimal probability density function of the random variable that should be added to the input is found to be a Dirac delta function, and hence the optimal random variable is a constant. The constant to be added depends upon the decision regions and the probability density functions under the two hypotheses, and is illustrated with an example. Also, an approximate procedure for the constant determination is derived for the mean-shifted binary hypothesis testing problem.

Here, we consider the problem of deciding between two hypotheses H_0 and H_1 that can occur with a priori probabilities $P[H_0] = \pi_0$ and $P[H_1] = \pi_1 = 1 = \pi_0$, respectively. Our criterion for performance will be probability of error P_e, although the derivation is easily modified to minimize the Bayes' risk by assigning costs associated with each decision [Kay 1998]. It is assumed that *the decision regions have already been specified, that they are not optimal in terms of minimizing* P_e, and that a single data sample x is used to make a decision. The already specified decision regions may be arbitrary and hence our solution encompasses such regions as if one would decide H_1 if $x > a$ or $|x| < a$ as examples. The single sample is usually a test statistic, i.e., a function of a set of observations, which is a common procedure in decision making. To improve the performance, "noisy sample" c is added to form $y = x + c$ prior to decision making. We allow c to be a *random variable* and determine the PDF of c that will yield the minimum P_e. It is proven next that the optimal PDF is a Dirac delta function, which leads to the conclusion that the *optimal random variable to be added is a degenerate one, i.e., a constant.*

4.1 Optimal PDF of Additive Noise Sample

To write the probability of error for the original problem, we define the decision rule (also called the test function or critical region indicator function) as

$$\phi(x) = \begin{cases} 0 & decide\ H_0 \\ 1 & decide\ H_1 \end{cases}.$$

Then, we have

$$\begin{aligned} P_e &= P[decide\ H_1 \mid H_0]P[H_0] + P[decide\ H_0 \mid H_1]P[H_1] \\ &= P[\phi(x) = 1 \mid H_0]\pi_0 + P[\phi(x) = 0 \mid H_1]\pi_1 \\ &= \pi_0 \int_{-\infty}^{\infty} \phi(x) p_0^X(x) dx + \pi_1 \int_{-\infty}^{\infty} (1 - \phi(x)) p_1^X(x) dx \end{aligned}$$

where $p_0^X(x)$ and $p_1^X(x)$ are the probability density functions (PDFs) under H_0 and H_1, respectively. This can be written as

$$P_e = \pi_1 + \int_{-\infty}^{\infty} \phi(x)\left[\pi_0 p_0^X(x) - \pi_1 p_1^X(x)\right] dx .$$

Now assume that we modify x by adding c so that the test stistic becomes $y = x + c$, where c is a random variable independent of x, and whose PDF is $p_C(c)$. Since the identical decision rule is to be used, we have

$$P_e = \pi_1 + \int_{-\infty}^{\infty} \phi(x)\left[\pi_0 p_0^Y(y) - \pi_1 p_1^Y(y)\right] dy .$$

But

$$p_0^Y(y) = \int_{-\infty}^{\infty} p_0^X(y-c) p_C(c) dc$$

$$p_1^Y(y) = \int_{-\infty}^{\infty} p_1^X(y-c) p_C(c) dc .$$

We have then that

$$\begin{aligned} P_e &= \pi_1 + \int_{-\infty}^{\infty} \phi(x)\left[\pi_0 \int_{-\infty}^{\infty} p_0^X(y-c) p_C(c) dc - \pi_1 \int_{-\infty}^{\infty} p_1^X(y-c) p_C(c) dc\right] dy \\ &= \pi_1 + \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} \phi(y)\left(\pi_0 p_0^X(y-c) - \pi_1 p_1^X(y-c)\right) dy\right] p_C(c) dc \\ &= \pi_1 + E_C\left[\int_{-\infty}^{\infty} \phi(y)\left(\pi_0 p_0^X(y-c) - \pi_1 p_1^X(y-c)\right) dy\right] \end{aligned}$$

where E_C denotes expected value with respect to $p_C(c)$. Hence, we wish to choose $p_C(c)$ so that the slightly more convenient form

$$J(p_C) = E_C\left[\int_{-\infty}^{\infty} \phi(y)\left(\pi_0 p_0^X(y-c) - \pi_1 p_1^X(y-c)\right) dy\right] \tag{4.1}$$

is *maximized.* This is done in the next section. We will see that the random variable C may be chosen as a constant and therefore we need only maximize the expression within the brackets of (4.1) over a constant c. But this is equivalent to shifting $\phi(u)$, the decision region function by $-c$. Hence, the solution effectively shifts the decision region by a constant. This suggests that another means for improving performance is to transform the decision region using a *nonlinear* transformation (instead of a simple shift). It can be done by transforming the data sample x using a nonlinear transformation g as $g(x)$. This is addressed in Chapter 8 and will be considered further in Phase II.

4.2 Derivation of Optimal PDF for *C*

It is well known that $E_C[g(C)]$ is maximized by placing all the probability mass at the value c for which $g(c)$ is maximized. We assume that the function $g(c)$ has at least one point at which a maximum is attained. Calling this point c_0 the optimal PDF is then $p_C(c) = \delta(c - c_0)$, where

$$c_0 = \arg_c \max g(c)$$

or

$$c_0 = \arg_c \max \int_{-\infty}^{\infty} \phi(y)\left(\pi_0 p_0^X(y-c) - \pi_1 p_1^X(y-c)\right)dy .$$

A slightly more convenient form for $g(c)$ is obtained by letting $u = y - c$ so that

$$g(c) = \int_{-\infty}^{\infty} \phi(u+c)\left(\pi_1 p_1^X(u) - \pi_0 p_0^X(u)\right)du \tag{4.2}$$

which is recognized as a *correlation* between $\phi(u)$ and $\pi_1 p_1^X(u) - \pi_0 p_0^X(u)$. In summary, *we should add the constant c to x, where c is the value that maximizes the correlation given in (2).* Since the decision function $\phi(x)$ in (4.2) is completely general, the optimal solution is valid for a given binary decision rule with *any decision region.* For example, if the original decision rule were to decide H_1 if $x > a$, then we would use $\phi(u) = 1$ for $u > a$, and zero otherwise in (4.2). If it were to decide H_1 if $|x| < a$, then we would use $\phi(u) = 1$ for $|u| < a$, and zero otherwise in (4.2). (Note that if $\phi(u) = 1$ for $\pi_1 p_1^X(u) - \pi_0 p_0^X(u) > 0$ and zero otherwise, then $g(c)$ is maximized for $c = 0$. This is because in this case the decision rule $\phi(u)$ is optimal.) In the next section we solve this for a given example.

4.3 The Gaussian Mixture Example

We now consider the problem described in [Kay 2000]. but instead choose the probability of error criterion. The problem is to decide between $p_0^X(x)$ and $p_1^X(x) = p_0^X(x-A)$, where $A > 0$ is a DC level that is known and the noise PDF is the Gaussian or normal mixture

$$p_0^X(x) = \tfrac{1}{2} N(x;\mu,\sigma^2) + \tfrac{1}{2} N(x;-\mu,\sigma^2) \tag{4.3}$$

where

$$N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right].$$

The original decision rule is to choose H_1 if $x > 0$ so that $\phi(x) = u_s(x)$, where $u_s(x)$ is the unit step function. Additionally, we assume equal a priori probabilities so that $\pi_0 = \pi_1 = ½$. As a result, we have from (4.2) that

$$\begin{aligned} g(c) &= \tfrac{1}{2}\int_{-\infty}^{\infty} u_s(u+c)\left(p_1^X(u) - p_0^X(u)\right)du \\ &= \tfrac{1}{2}\int_{-c}^{\infty} \left(p_1^X(u) - p_0^X(u)\right)du \\ &= \tfrac{1}{2}[(1 - F_1(-c)) - (1 - F_0(-c))] \\ &= \tfrac{1}{2}[F_0(-c) - F_1(-c)] \end{aligned}$$

where F_i is the cumulative distribution function of x under the hypothesis H_i. For our problem, we have that $p_1^X(x) = p_0^X(x-A)$ and so $F_1(x) = F_0(x-A)$. Thus,

$$g(c) = \tfrac{1}{2}[F_0(-c) - F_0(-c - A)]$$

and differentiating and setting equal to zero produces

$$p_0^X(-c) = p_0^X(-c - A)$$

or equivalently since $p_0^X(x)$ is even, we have the general requirement

$$p_0^X(c) = p_0^X(c + A)\,. \tag{4.4}$$

Using (4.4) in (4.3) produces

$$\phi(c;\mu,\sigma^2) + \phi(c;-\mu,\sigma^2) = \phi(c + A;\mu,\sigma^2) + \phi(c + A;-\mu,\sigma^2)$$

which upon simplification yields the equation

$$\exp(\mu c/\sigma^2) + \exp(-\mu c/\sigma^2) = \exp\left[\left(-c(A-\mu)/\sigma^2\right) - A^2/(2\sigma^2) + \mu A/\sigma^2\right] + \exp\left[\left(-c(A+\mu)/\sigma^2\right) - A^2/(2\sigma^2) - \mu A/\sigma^2\right].$$

For $\mu = 3$, $\sigma^2 = 1$, and $A = 1$, we have

$$\exp(3c) + \exp(-3c) = \exp(2c + 5/2) + \exp(-4c - 7/2)\,.$$

The exact value of c found through a numerical search is $c = 2.50$, which could also be found by ignoring the terms exp(-3c) and exp(-4c – 7/2) since these are nearly zero for this value of c. Another solution is found by ignoring the other set of terms to yield $c = -3.5$. Note that either of these choices causes the PDFs of $x + c$ under H_0 and H_1 to cross at the origin (see Figs. 4.1 and 4.2). If we did not have the right-most Gaussian mode, then the choice of $c = 2.5$ would result in a maximum likelihood (ML) receiver, which is optimum [Kay 1998]. This is because a maximum likelihood receiver chooses the hypothesis whose PDF value is larger. In our case, the *fixed* decision regions are $R_1 = \{x: x > 0\}$ for H_1 and $R_0 = \{x: x < 0\}$ for H_0 as shown in Fig.4.1. These decision regions are not optimal. The optimal ML decision regions are indicated in Fig.4.1 as R_0^* and R_1^*. Therefore, the region in x for which $R_0^* \neq R_1^*$, which corresponds to the dark PDF lines, will result in incorrect decisions. By the addition of c, however, the extent of this incorrect decision region is reduced, as indicated in Fig.4.2.

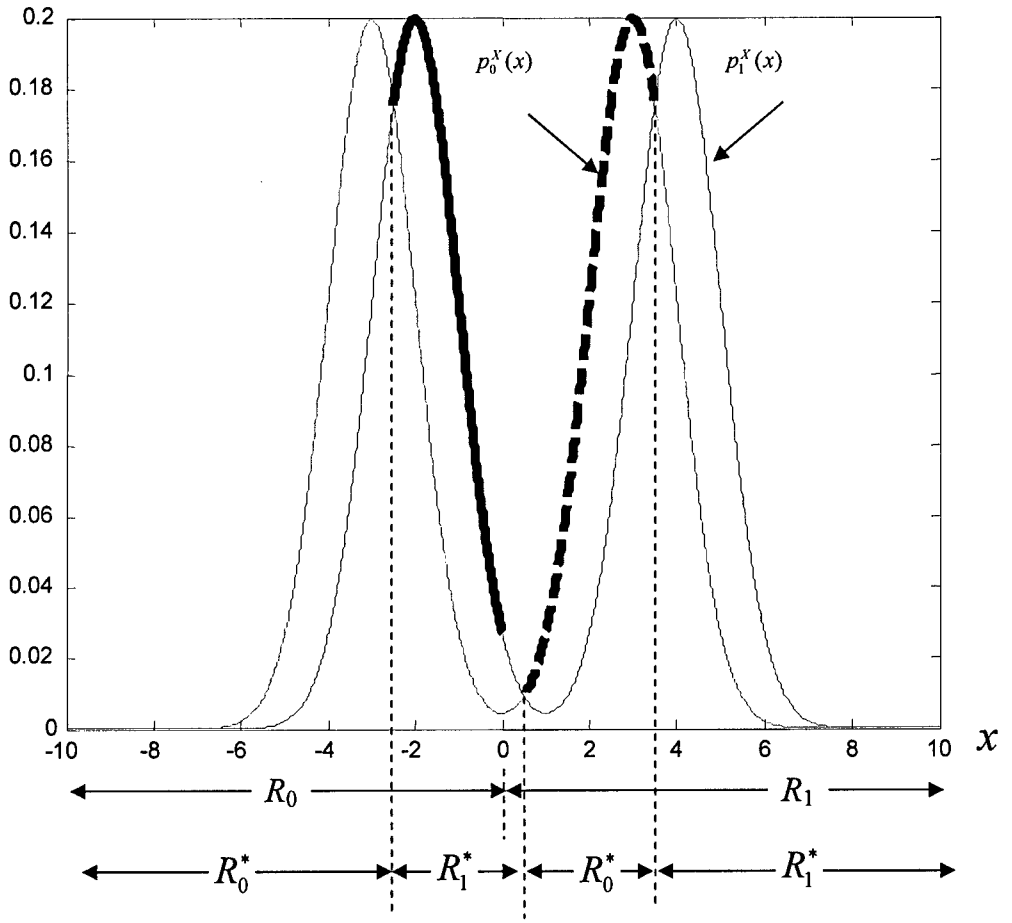


Fig.4.1 Original PDFs. The left-most PDF modes cross at x = -2.5, which is indicated by the dashed vertical line. The *fixed* decision regions are indicated by R_i while the optimal ML decision regions are indicated by R_i^*.

It is instructive to also plot the probability of error versus c or equivalently the probability of correct decision $P_c = 1 - P_e$ versus c. This is shown in Fig.4.3. Note that as expected the probability of a correct decision is maximized at c = 2.5 and also at c = -3.5. This type of curve is normally associated with the phenomenon of stochastic resonance, although here we see that it is *not unimodal*. This result is unlike that reported in [1-3] and so debunks the common assumption that adding too much noise will degrade performance. The latter is only true if the performance is unimodal.

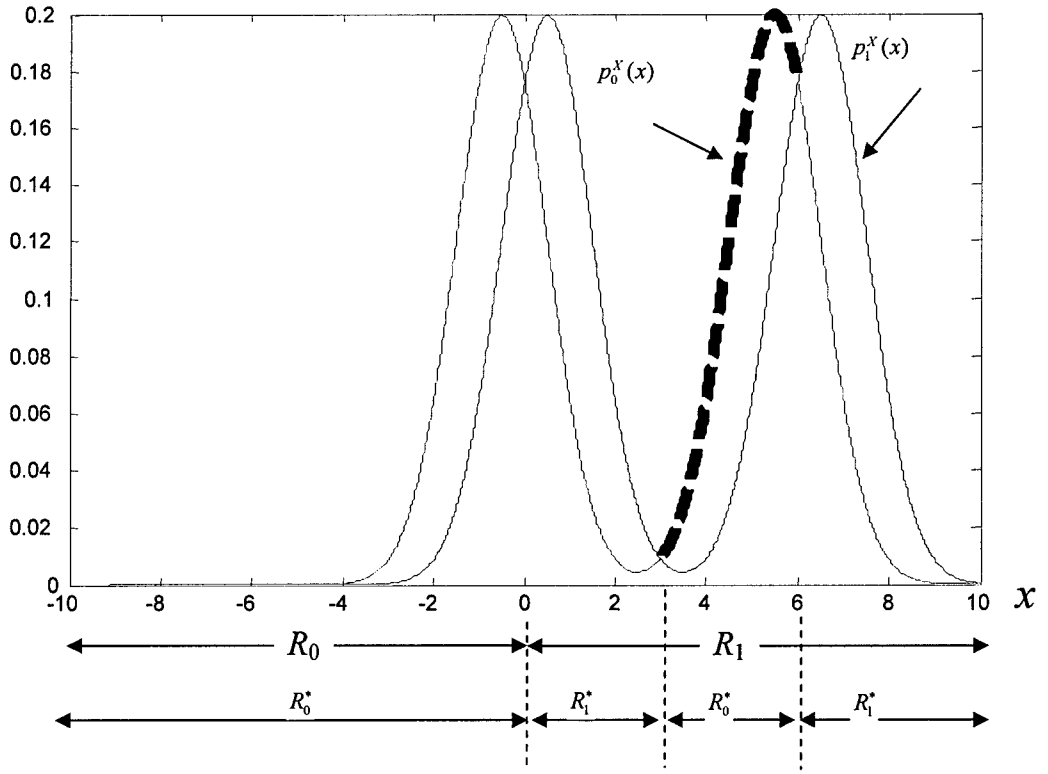


Fig.4.2 PDFs after $c = 2.5$ is added to x. The *fixed* decision regions are indicated by R_i while the optimal ML decision regions are indicated by R_i^*.

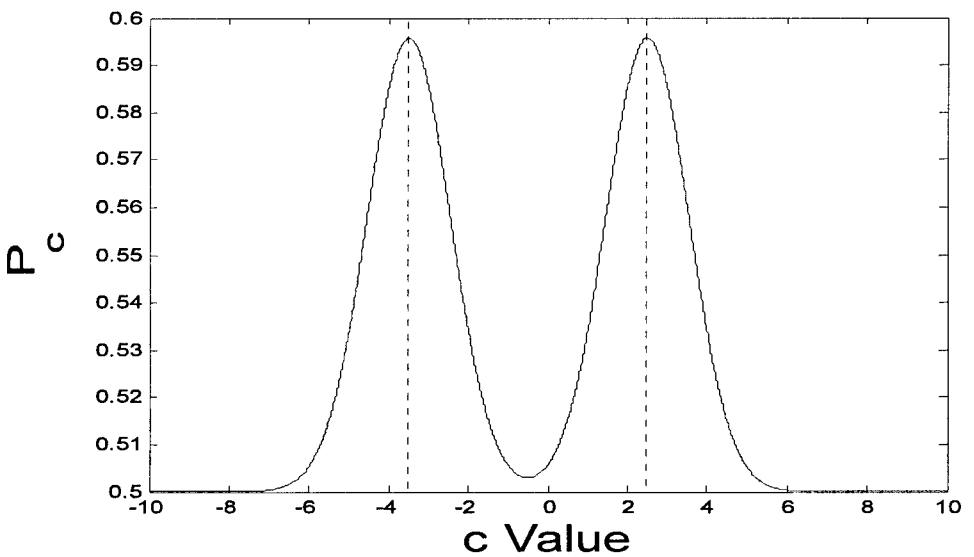


Figure 4.3 Probability of a correct decision versus the value of the constant c to be added to the data sample. The dashed lines are at $c = -3.5$ and $c = 2.5$.

5.0 Application of Stochastic Resonance to Nonparametric Detectors

Here we consider detection performance of two additional nonparametric detectors which exhibit improvement via additive SR noise; namely, the Wilcoxon and the dead-zone limiter detectors. In addition to the sign detector, these detectors were considered in [Chen, et. al., 2006c]. The asymptotic efficiency (AE) as well as finite sample detection performance of these SR modified detectors was reported. For large sample sizes, the AE of the Wilcoxon and the dead-zone limiter [Kassam 1976] detectors was shown not to improve by the addition of SR noise. However, for finite sample sizes, both of these detectors show improvement in the presence of additive SR noise.

Nonparametric detectors have received considerable attention in signal detection problems [Kassam 1980]. An important feature of such detectors is their guaranteed *level and reasonable power* for large classes of input distributions. However, in most cases, a nonparametric detector is less efficient than the optimal detector. Therefore, an important consideration is the potential improvement of their performance while maintaining their false alarm rate (CFAR) property. Here, we explore the potential detection performance improvement of several nonparametric detectors by adding SR noise to the observed data.

5.1 Problem Formulation for Nonparametric Detectors

Let us consider a detection problem based on the observed data vector $\mathbf{x} = [x_1, x_2, \ldots, x_N]$ with probability density function $p(\mathbf{x})$, where the x_i, $i = 1, 2, \ldots, N$ are independent identically distributed (i.i.d.) scalar random variables. We decide between hypotheses H_1 and H_0 given by

$$\begin{cases} H_0 : p(\mathbf{x}) = \prod_{i=1}^{N} f_x(x_i) \\ H_1 : p(\mathbf{x}) = \prod_{i=1}^{N} f_x(x_i - A) \end{cases} \tag{5.1}$$

where the pdf $f_x(\cdot)$ of the scalar random variable x_i is symmetric, i.e., $f_x(x) = f_x(-x)$ and $A > 0$. Therefore, this test is essentially the detection of a constant positive DC signal A in additive noise. with a symmetric pdf.

Following the SR approach, detection performance enhancement is achieved by adding noise $\mathbf{n} = [n_1, n_2, \ldots, n_N]$ to the original data process $\mathbf{x}$ to obtain a new vector $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where n_i are i.i.d. scalar random variables with pdf $f_n(n)$. The constant false alarm rate (CFAR) property is maintained by retaining the symmetric pdf property of $\mathbf{x}$. Therefore, we restrict $f_n(n)$ to be symmetric, i.e., $f_n(n) = f_n(-n)$. Given $f_x(x)$ and $f_n(n)$, the pdf of y_i under the H_0 hypothesis can be expressed by the convolutions of the pdfs such that

$$\begin{aligned} f_y(y_i) &= f_x(x_i) * f_n(n_i) \\ &= \int_{-\infty}^{\infty} f_x(x_i) f_n(y_i - n_i) dx_i \end{aligned}$$

$$= \int_{-\infty}^{\infty} f_x(y_i - n_i) f_n(x_i) dx_i \,. \tag{5.2}$$

It can be shown that $f_y(y) = f_y(-y)$, i.e., $f_y(y)$ is a symmetric function. Therefore, $P(y > 0|H_0) = 1/2.$, and the CFAR property of the nonparametric detectors is maintained.

The binary hypotheses testing problem for this new observed data **y** can be expressed as:

$$\begin{cases} H_0 : p(\mathbf{y}) = \prod_{i=1}^{N} f_y(y_i) \\ H_1 : p(\mathbf{y}) = \prod_{i=1}^{N} f_y(y_i - A) \end{cases} . \tag{5.3}$$

The cumulative distribution function (cdf) of y_i is given by

$$\begin{aligned} F_y(y_i) &= \int_{-\infty}^{y_i} \int_{-\infty}^{\infty} f_x(x_i) f_n(y_i - x_i) dx_i dy_i \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{y_i} f_x(x_i) f_n(y_i - x_i) dx_i dy_i \\ &= \int_{-\infty}^{\infty} f_x(x_i) F_n(y_i - x_i) dx_i = \int_{-\infty}^{\infty} f_n(x_i) F_x(y_i - x_i) dx_i \,. \end{aligned} \tag{5.4}$$

5.2 SR Detection Performance for Nonparametric Detectors

Detection performance for the three detectors is now considered using stochastic resonance for the problem involving a known DC level in Gaussian mixture noise with mean $\mu = 3$ and modal variance $\sigma^2 = 1$. In the asymptotic case where signal strength vanishes and sample sizes approach infinity, the performance is evaluated in terms of the asymptotic relative efficiency (ARE) between the original detector and the SR noise modified detector. Further, the ARE can often be expressed as the ratio of their asymptotic efficiencies given by

$$E = \lim_{N \to \infty} \frac{\left\{ \frac{d}{dA} E[T(\mathbf{x}_N)] |_{A=0} \right\}^2}{N Var_{A=0}[T(\mathbf{x}_N)]} \tag{5.5}$$

where $T(\cdot)$ is the test statistic. Similarly, for the finite sample case, we compare the relative performances by the deflection measure [Picinbono 1995] which is defined as

$$D(T) = \frac{[E(T \mid H_0) - E(T \mid H_1)]^2}{\mathrm{var}(T \mid H_0)} \tag{5.6}$$

for the sign and Wilcoxon detectors. For the dead-zone limiter detector, we illustrate the performance by several examples.

A. The Sign Detector

For the sign detector, we have a test statistic and decision rule as follows

$$T_s = \sum \mathrm{sgn}(\mathrm{x_i}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{5.7}$$

where sgn(x) is the sign of x, given by

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases}. \tag{5.8}$$

Let $P_i^{\mathrm{x}} = \mathrm{P}(\mathrm{sgn}(\mathrm{x}) = 1|\mathrm{H_i})$, i = 0, 1. From, (5.1), we have $P_0^{\mathrm{x}} = \int_0^\infty f(x)dx = 0.5$ and

$$\begin{aligned} P_1^{\mathrm{x}} &= \int_0^\infty f(x-A)dx \\ &= \int_{\mathrm{A}}^\infty f(x)dx = F_{\mathrm{x}}(\mathrm{A}) = \frac{1}{2}\left[Q\left(\frac{\mu - A}{\sigma_0^2}\right) + Q\left(\frac{-\mu - A}{\sigma_0^2}\right)\right] > 0.5. \end{aligned} \tag{5.9}$$

Furthermore, the test statistic $\mathrm{T_s}$ is binomially distributed with parameter $\mathrm{P_i}$ under $\mathrm{H_i}$, i = 0, 1. Since, $P_0^{\mathrm{x}} = 0.5$ is fixed, therefore when $P_1^{\mathrm{x}} > 0.5$, the detection performance of the sign detector is monotonically determined by P_1^{x}, i.e., the higher the value of P_1^{x}, the better the detection performance of (5.7). It can also be shown that the expected value of $\mathrm{T_s}$ under $\mathrm{H_i}$ is expressed as $\mathrm{E(T_s|H_i)} = \mathrm{N}P_i^{\mathrm{x}}$, i = 0, 1 and the variance of T under $\mathrm{H_0}$ is $\mathrm{var(T|H_0)} = \mathrm{N/4}$. The deflection measure of the sign detector $\mathrm{D_s}$ is given as

$$D_s^{\mathrm{x}} = 4N(P_1^{\mathrm{x}} - P_0^{\mathrm{x}})^2 = 4N(P_1^{\mathrm{x}} - 0.5)^2. \tag{5.10}$$

Similarly, for y, we have $P_0^{\mathrm{y}} = 0.5 = P_0^{\mathrm{x}}$ and $D_s^{\mathrm{y}} = 4N(P_1^{\mathrm{y}} - 0.5)^2$. However, due to the additive SR noise n, P_1^{y} is changed such that

$$\begin{aligned} P_1^{\mathrm{y}} &= \int_{\mathrm{A}}^\infty f(y)dy \\ &= \int_{\mathrm{A}}^\infty \int_{-\infty}^\infty f_x(y-x) f_n(x)dxdy \\ &= \int_{-\infty}^\infty f_n(x) \int_{A-x}^\infty f_x(y)dydx \\ &= \int_{-\infty}^\infty f_n(x) F_x(A+x)dx \\ &= \int_{-\infty}^\infty \frac{1}{2} f_n(x)[F_x(A+x) + F_x(A-x)]dx \end{aligned}$$

$$= \int_{-\infty}^{\infty} f_n(x)G(x)dx \tag{5.11}$$

The right hand side of (5.11) is obtained by applying the symmetry property of f_n(x) and letting G(x) = $\frac{F_x(A+x)+F_x(A-x)}{2}$. From (9), P_1^x = G(0). Let G_M = max(G(x)) and x_0^g be the minimum non-negative x such that G(x_0^g) = G_M. Since $\int_{-\infty}^{\infty} f_n(x)dx = 1$, we have $P_1^y \le G_M$. The equality can be obtained be selecting an optimum SR noise pdf f_n^0 such that

$$f_n^0(x) = \frac{1}{2}\delta(x - x_0^g) + \frac{1}{2}\delta(x + x_0^g). \tag{5.12}$$

Therefore, when G_M > G(0), the detection performance of the sign detector can be improved by adding SR noise. Correspondingly, its deflection measure is such that $D_s^{y_o} > D_s^x$. Note that G(x) can be expressed as

$$\begin{aligned} G(x) &= \frac{F_x(A + x) + F_x(A - x)}{2} \\ &= \frac{F_x(A + x) + 1 - F_x(-A + x)}{2} \\ &= \frac{1}{2} + \frac{1}{2}\int_{A+x}^{A+x} f_x(t)dt. \end{aligned} \tag{5.13}$$

Therefore, for the asymptotic case where N $\to \infty$ and A $\to$ 0, it follows that G(x) $\approx$ ½ + Af_x(x) so that G(0) $\approx$ ½ + Af_x(0). As a result, the asymptotic detection performance can be improved if $f_x(0) \neq \max(f_x(x))$. The same conclusion can also be obtained by evaluating the asymptotic efficiency and the ARE between the original detector and the SR noise enhanced detector. For the sign detector, its asymptotic efficiency is given by

$$E_s^x = 4f_x^2(0) \tag{5.14}$$

and similarly

$$E_s^y = 4f_y^2(0) = 4\left[\int f_x(x)f_n(x)dx\right]^2. \tag{5.15}$$

Again since $\int_{-\infty}^{\infty} f_n(x)dx = 1$, the optimum SR pdf for this case is

$$f_n^0(x) = \frac{1}{2}\delta(x - x_0) + \frac{1}{2}\delta(x + x_0) \tag{5.16}$$

where $f_x(x_0) = \max(f_x(x))$. In general, we have the ARE between the noise modified detectors and the original detectors $E_s^{y,x}$ given by

$$E_s^{y,x} = \frac{f_y^2(0)}{f_x^2(0)}. \tag{5.17}$$

B. The Wilcoxon Detector

The Wicoxon detector test statistic is expressed as

$$T_W = \sum_{j=1}^{N} \sum_{i=1}^{i=j} \mathrm{sgn}(x_i + x_j) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{5.18}$$

In the asymptotic case, the asymptotic efficiency of the original Wilcoxon detector E_W^x is given by $E_W^x = 12\left[\int_{-\infty}^{\infty} f_x^2(x)dx\right]^2$. Let $H_x(\omega) = \int_{-\infty}^{\infty} f_x(x)\exp(-j\omega x)dx$ be the Fourier transform of $f_x(\cdot)$. Since $f_x(x) \geq 0$ is a symmetric real function, $H_x(\omega) = \int_{-\infty}^{\infty} f_x(x)\cos(\omega x)dx$ is also a real function. From Parseval's theorem, we have

$$E_W^x = 12\left[\int_{-\infty}^{\infty} f_x^2(x)dx\right]^2 = \frac{6}{\pi}\left[\int_{-\infty}^{\infty} H_x^2(\omega)d\omega\right]^2. \tag{5.19}$$

Similarly, the asymptotic efficiency of the SR noise modified detector E_W^y can be expressed as

$$E_W^y = 12\left[\int_{-\infty}^{\infty} f_y^2(y)dy\right]^2 = \frac{6}{\pi}\left[\int_{-\infty}^{\infty} H_y^2(\omega)d\omega\right]^2. \tag{5.20a}$$

From (5.2), we have $H_y(\omega) = H_x(\omega)\, H_n(\omega)$, so that

$$E_W^y = \frac{6}{\pi}\left[\int_{-\infty}^{\infty} H_y^2(\omega)H_n^2(\omega)d\omega\right]^2 \tag{5.20b}$$

and the ARE between the noise modified detectors and the original detectors is given by

$$E_W^{y,x} = \frac{\left(\int_{-\infty}^{\infty} f_y^2(y)dy\right)^2}{\left(\int_{-\infty}^{\infty} f_x^2(x)dx\right)^2}. \tag{5.21}$$

Note that

$$
\begin{aligned}
|H_n(\omega)| &= \left| \int_{-\infty}^{\infty} f_n(x) \exp(-j\omega x) dx \right| \\
&= \left| \int_{-\infty}^{\infty} f_n(x) \cos(\omega x) dx \right| \\
&\le \int_{-\infty}^{\infty} f_n(x) |\cos(\omega x)| \, dx \\
&\le \int_{-\infty}^{\infty} f_n(x) dx = 1.
\end{aligned}
$$

Therefore, we have $H_y^2(\omega) \le H_x^2(\omega)$ and furthermore, $E_W^y \le E_W^x$. In other words, in the asymptotic limit where N $\rightarrow \infty$ and A $\rightarrow$ 0, the Wilcoxon detector performance cannot be improved by adding independent SR noise. However, the detection performance may still be improved in the finite sample case.

C. The Dead-Zone Limiter Detector

The dead-zone limiter detector [Kassam 1976] employes the dead-zone limiter characteristic l_c to operate on the data where $l_c(\cdot)$ is given by

$$
l_c(x) = \begin{cases} 1 & x > c \\ 0 & -c \le x \le c \\ -1 & x < c \end{cases} \tag{5.22}
$$

where c is a prescribed positive number. Let N_{cp} be the number of samples which satisfy $x_i > c$ and N_c be the number of samples which satisfy $|x_i| > c$. In order to obtain a false alarm rate, α, the dead-zone limiter detector selects the H_1 hypothesis with probability one when $N_{cp} > g_\alpha(N_c)$ and with probability $\beta_\alpha(N_c)$ when $N_{cp} = g_\alpha(N_c)$. Both $g_\alpha(N_c)$ and $\beta_\alpha(N_c)$ are suitable functions such that the false alarm rate is fixed at α. For the dead-zone limiter detector with parameter c, assuming $F_x(c) < 1$ and $f_x(c)$ is continuous at c, we have its asymptotic efficiency E_{DZ}^x given by

$$
E_{DZ}^x = 2 \frac{f_x^2(c)}{1 - F_x(c)}. \tag{5.23a}
$$

Thus, for the SR noise modified detector, its asymptotic efficiency is expressed as

$$
E_{DZ}^y = 2 \frac{f_y^2(c)}{1 - F_y(c)}. \tag{5.23b}
$$

The ARE between the SR noise modified dead-zone limiter detector and the original dead-zone limiter detector is given by

$$E_{DZ}^{y,x} = \frac{f_y^2(c)/(1-F_y(c))}{f_x^2(c)/(1-F_x(c))}. \tag{5.24}$$

Using (5.2) in (5.23b), we have

$$\begin{aligned} E_{DZ}^{y} &= 2\frac{f_y^2(c)}{1-F_y(c)} \\ &= \frac{2\left(\int f_x(c-x)f_n(x)dx\right)^2}{1-\int F_x(c-x)f_n(x)dx} \\ &= \frac{2\left(\int f_x(c-x)f_n(x)dx\right)^2}{\int(1-F_x(c-x))f_n(x)dx} \\ &= \frac{2\left(\int f_x^2(c-x)f_n(x)dx\right)}{\int(1-F_x(c-x))f_n(x)dx}. \end{aligned} \tag{5.25}$$

Let $\kappa = \max \frac{2f_x^2(c-x)}{1-F_x(c-x)}$. Therefore, when c is selected to maximize E_{DZ}^{x}, i.e., $\kappa = \frac{2f_x^2(c)}{1-F_x(c)}$, we have $E_{DZ}^{y} \leq E_{DZ}^{x}$. Thus, the asymptotic efficiency of the tuned dead-zone limiter detector with optimal parameter c cannot be improved by adding SR noise. However, (5.25) does not rule out the possible SR effect when c is not optimum. Furthermore, similar to the Wicoxon detector, for the finite sample and vanishing signal case, the detection performance of the dead-zone limiter detector may still be improved by adding suitable SR noise.

5.3 Experimental Results

Here we consider the detection of a known DC signal in symmetric Gaussian mixture noise; i.e.,

$$f_x(x) = \tfrac{1}{2}\gamma(x;-\mu,\sigma_0^2) + \tfrac{1}{2}\gamma(x;\mu,\sigma_0^2)$$

where

$$\gamma(x;\mu,\sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

is the PDF of a Gaussian random variable with mean μ and variance σ^2. We consider two types of SR noise. These include the symmetric two-peak random noise with two random values

$f_s(x) = 0.5\delta(x - \tau) + 0.5\delta(x + \tau)$ *Symmetric two-peak SR noise*

and white Gaussian SR noise

$f_g(x) = \gamma(x;0,\tau^2)$. *Gaussian SR noise*

The noise modified data processes are denoted as y_s and y_g, respectively. In this example, we set $\sigma_0 = 1$, $\mu = 3$, and the sample size N = 5. From (5.2) we have the PDF of y_g given by

$$f_{y_g}(y) = \tfrac{1}{2}\gamma(y; -\mu, \sigma_0^2 + \tau^2) + \tfrac{1}{2}\gamma(y; \mu, \sigma_0^2 + \tau^2) \tag{5.26}$$

and the PDF of y_s

$$f_{y_s}(y) = \tfrac{1}{4}\gamma(y; -\mu - \tau, \sigma_0^2) + \tfrac{1}{4}\gamma(y; -\mu + \tau, \sigma_0^2) + \tfrac{1}{4}\gamma(y; \mu - \tau, \sigma_0^2) + \tfrac{1}{4}\gamma(y; \mu + \tau, \sigma_0^2). \tag{5.27}$$

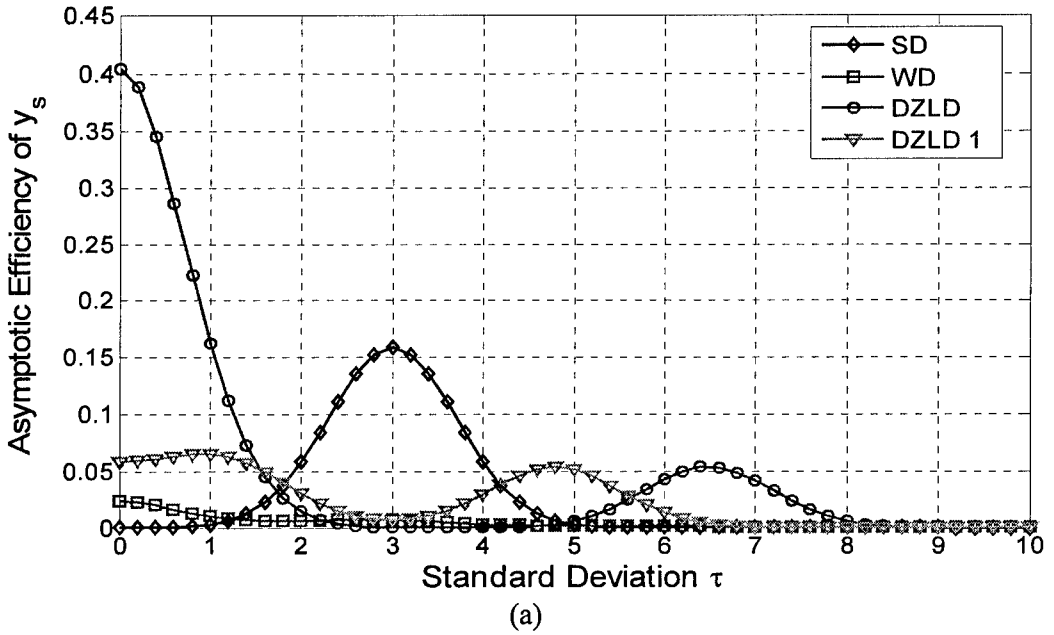
Next, we evaluate both the asymptotic detection performance and the finite sample detection performance for this particular detection problem for the three nonparametric detectors.

A. Asymptotic Detection Performance

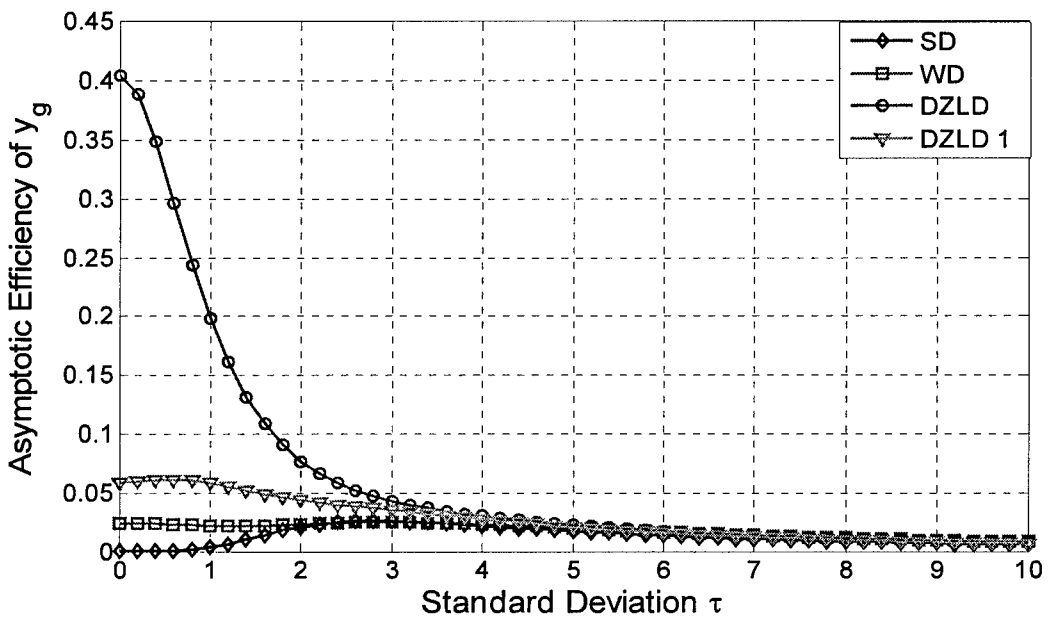
In this section, the asymptotic efficiency of the three SR modified nonparametric detectors foe both SR noises are obtained and plotted in Fig. 5.1. The detection performance of nonparametric detectors based on y_s and y_g are denoted as E^{y_s} and E^{y_g}, respectively. For the dead-zone limiter detector, two different c values are examined. One is the optimum value of the dead-zone limiter $c_0 = 3.61$ for the problem considered here and its corresponding values are shown as $E_{DZ}^{y_s}$ and $E_{DZ}^{y_g}$ in the figure. The other parameter is $c_1 = 0.61\sqrt{\mu^2 + \sigma_0^2} = 1.929$ which is the optimum value assuming that the noise is Gaussian distributed with the same variance as that in our example. In both Fig.5.1a and Fig.5.1b, the asymptotic efficiency of the Wilcoxon detector and the dead-zone limiter with $c_0 = 3.61$ is maximum when $\tau = 0$, i.e., in the limit of large data samples, the detection performance of the optimal dead-zone limiter and the Wilcox detector cannot be improved by adding SR noise. However, for the sign detector and the suboptimal dead-zone limiter ($c_1 = 1.929$), their detection performance can be enhanced! For the sign detector based on y_s, from (5.16) we have $f_x(\tau_0) = \max(f_x(x))$. Since f_x is a symmetric Gaussian mixture noise and $2\mu = 6\sigma_0$, the distance between the two peaks is significantly larger than their variances and the maximum value of f_x is reached at the mean value of each component of the mixture, i.e., $\tau_s = \mu = 3$. Thus, we have the maximum achievable asymptotic efficiency $E_S^{y_s} = 0.1592$. Compared to the original Sign detector which has a low value of asymptotic efficiency $E_S^x = 7.8565\text{x}10^{-5}$, the ARE between these two detectors is $E_S^{y_s^0, x} = 2026$, i.e., by adding suitable SR noise, the detection performance of the sign detector is enhanced by more than a factor of 2000. Similarly, for the Gaussian SR noise case, we have the optimum $\tau_g = \sqrt{\mu^2 - \sigma_0^2} = 2.8284$. Furthermore, the maximum achievable asymptotic efficiency when SR Gaussian noise is added is $E_S^{y_g^0} = 0.026$ and the corresponding ARE is $E_S^{y_g^0, x} = 331$.

For the Wilcoxon detector, the results shown in Fig.5.1 are to be expected. When $\tau > 0$, the efficiencies $E_W^{y_s}$ and $E_W^{y_g}$ are always less than E_W^x, i.e., adding any noise to the observation data will only degrade the detection performance. The same conclusion can be drawn for the dead-zone limiter detector with the optimal parameter $c_0 = 3.61$. However, for the dead-zone limiter detector with suboptimal $c_1 = 1.929$, as shown in Fig.5.1a and Fig.5.1b, when τ is relatively

small, $E_{DZ,1}$ increases when τ increases, i.e., the detection performance is improved by adding suitable SR noise. For y_s, the maximum $E_{DZ,1}^{y_s}$ is $E_{DZ,1}^{y_s} = 0.0657$ with parameter $\tau_{DZ}^{y_s} = 0.92$ and the maximum $E_{DZ,1}^{y_g} = 0.0614$ is achieved with parameter $\tau_{DZ}^{y_g} = 0.590$ for y_g. The ARE for both cases are $E_{DZ,1}^{y_s,x} = 1.116$ and $E_{DZ,1}^{y_g,x} = 1.0416$ for y_s and y_g, respectively.



(a)



(b)

Fig. 5.1 Asymptotic Efficiency of the SR Noise Modified Nonparametric Detectors, (a) based on y_s, (b) based on y_g.

B. Finite Sample Size Detection Performance

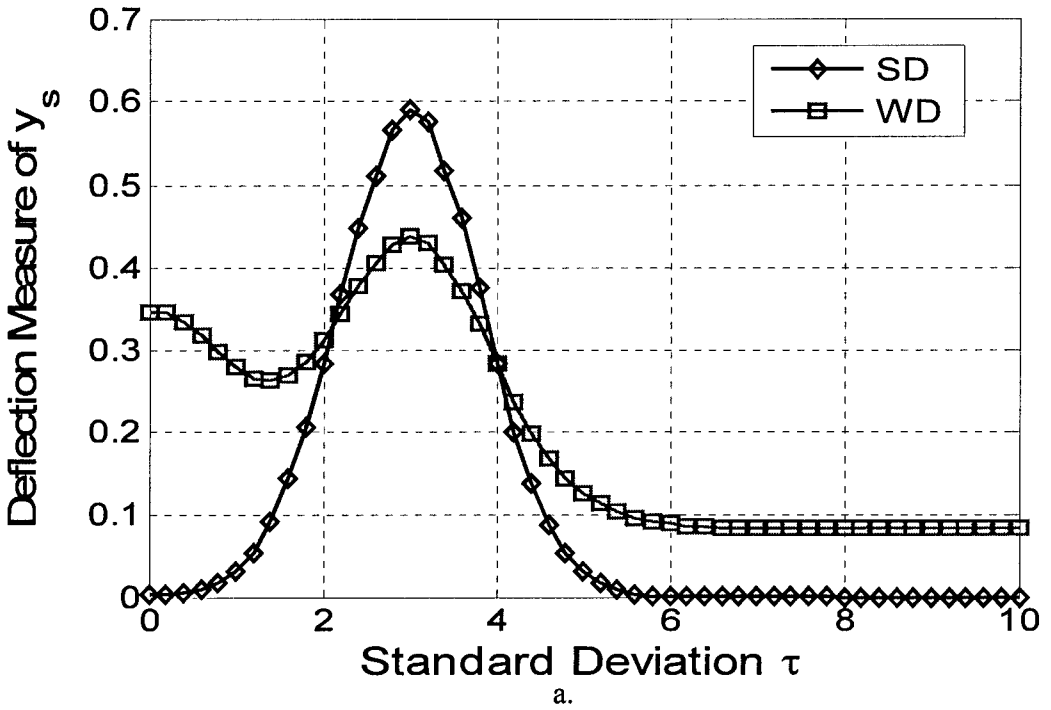
In this section, we consider two different values of A to examine the detection performance of the nonparametric detectors. For the sign and Wilcoxon detector, A = 1 is assumed in the first example. From (5.9), we have

$$P_1^x = F_x(A) = \frac{1}{2}Q\left(\frac{\mu - A}{\sigma_0}\right) + \frac{1}{2}Q\left(\frac{-\mu - A}{\sigma_0}\right)$$

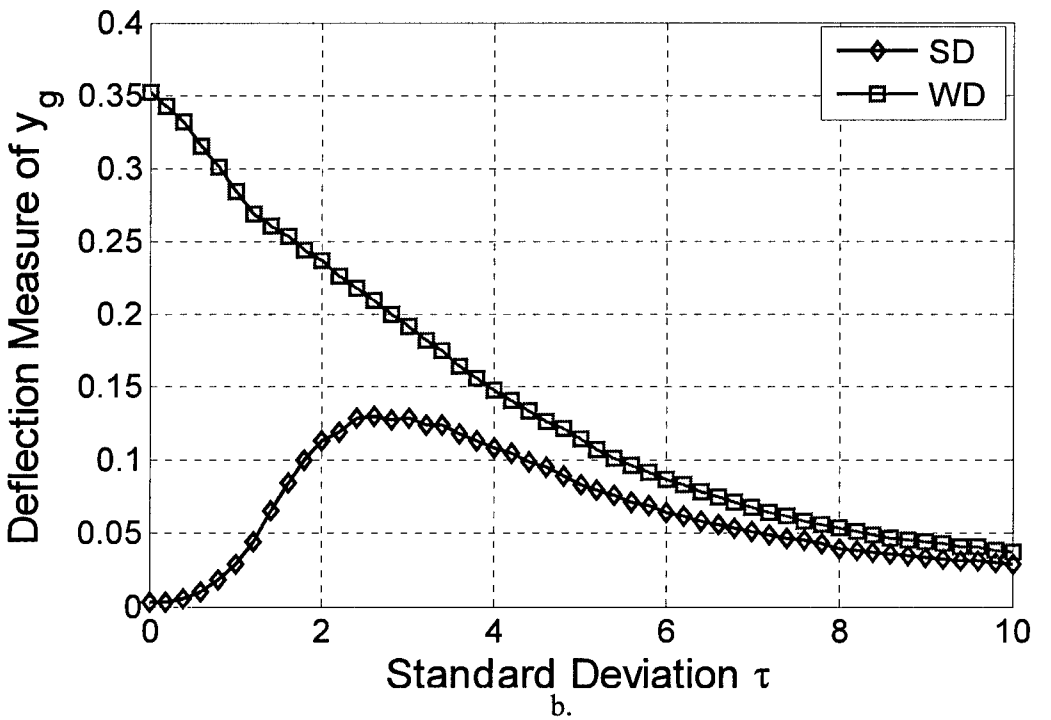
where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right] dt$ is the complementary distribution function for the standard Gaussian distribution. Therefore, we have in this case,

$$\begin{aligned} G_x(x) &= \frac{F_x(A+x) + F_x(A-x)}{2} \\ &= \frac{1}{4}Q\left(\frac{\mu - A - x}{\sigma_0}\right) + \frac{1}{4}Q\left(\frac{\mu - A + x}{\sigma_0}\right) + \frac{1}{4}Q\left(\frac{-\mu - A - x}{\sigma_0}\right) + \frac{1}{4}Q\left(\frac{-\mu - A + x}{\sigma_0}\right). \end{aligned} \quad (5.28)$$

Taking the derivative w.r.t. x, setting it equal to zero and solving, we have $x_0 \approx \mu = 3$. The optimal SR noise PDF for the sign detector is $f_n^o(x) = \frac{1}{2}\delta(x+3) + \frac{1}{2}\delta(x-3)$. Thus, we have $P_1^{y_o}$ = 0.6707 and the maximum deflection measure $D_S^{y_o}$ = 0.5826. The relationships between the deflection measure D and τ for the sign and Wilcoxon detectors obtained by Monte Carlo simulation are shown in Fig.5.2. Clearly both curves achieve their peak values when $\tau > 0$, i.e., an improvement in detection performance is obtained when suitable noise is added. The deflection measure of the SR noise modified sign detectors for both types of SR noises improves when the SR noise is suitable. On the other hand, for the Wilcoxon detector, it is a different story. As shown in Fig.5.2a, when the SR noise is a randomization of two values, the deflection measure $D_W^{y_s}$ may be larger than that of the original detector as noted by the peak at τ = 3. However, this is not the case when the additive SR noise is Gaussian. As shown in Fig.5.2b, $D_W^{y_s}$ decreases monotonically as τ increases, i.e., for the Wilcoxon detector, the SR phenomenon is not observed for Gaussian SR noise.



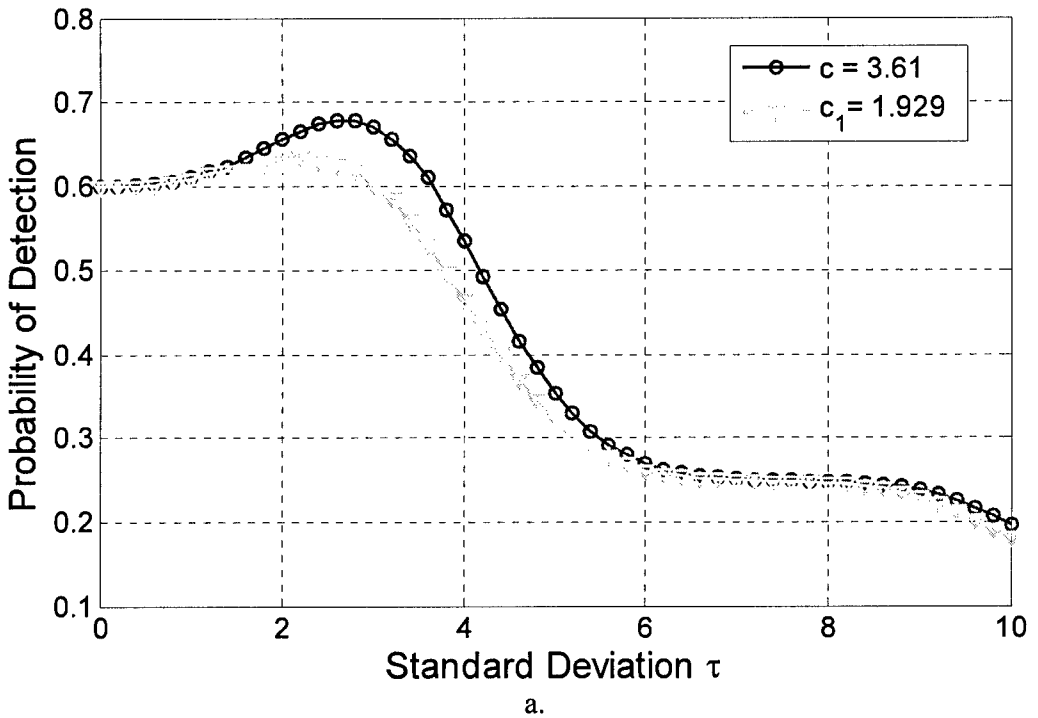
a.



b.

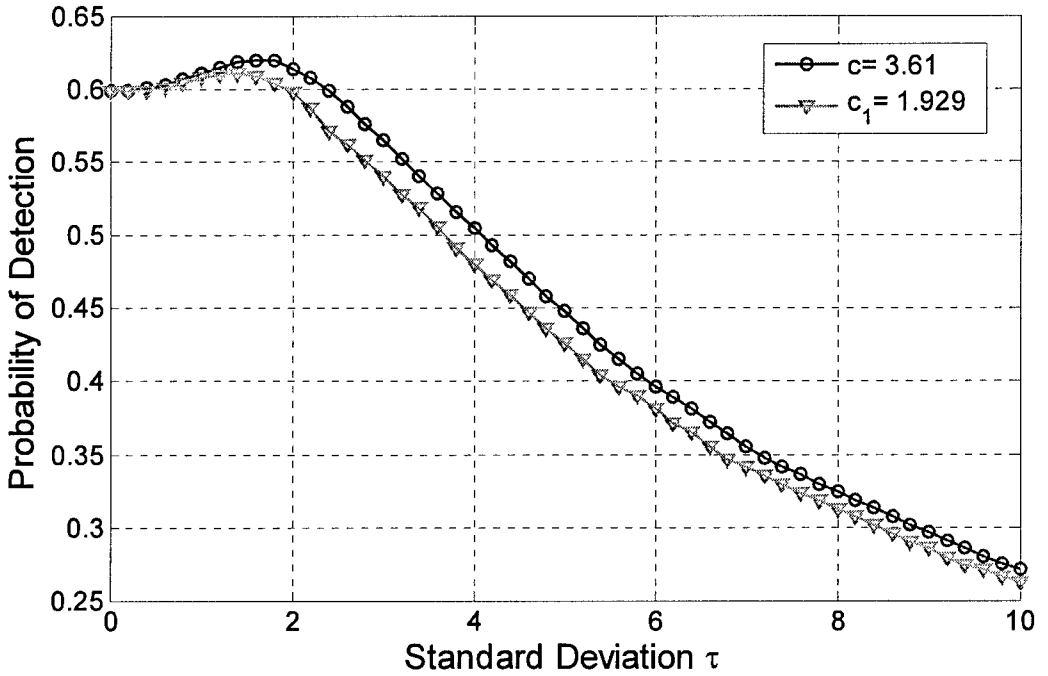
Fig. 5.2 Detection performance for the sign detector and the Wilcoxon detector using a finite sample size N = 5 and signal strength A = 1; a.) based on y_s b.) based on y_g.

For the dead-zone limiter detector, we chose a relatively strong signal with A = 4. Again, two dead-zone limiter detectors with parameters $c_0 = 3.61$ and $c_1 = 1.929$ are tested. Since the dead-zone detector is a conditional detector, unlike the other two detectors, it is difficult to calculate its deflection measure. To evaluate its relative detection performance, we use the detection power β of the SR noise enhanced dead-zone limiter while keeping the false alarm rate $\alpha = 0.1$ fixed. Intensive Monte Carlo simulations were performed to obtain the detection performance. As shown in Fig.5.3a, for both dead-zone limiter detectors based on observation y_s, the detection performance increases as $\tau > 0$ increases and reach their peaks when $\tau \approx 2.8$ and $\tau \approx 2.1$ for c_0 and c_1, respectively. When the additive SR noise is Gaussian, maximum detection powers β are obtained when $\tau \approx 1.8$ and $\tau \approx 1.4$, respectively. Also, by comparing the maximum values from these two figures, we find that the detection performance of y_s is better than that of y_g, i.e., in this case, adding two peak random SR noise is better than adding Gaussian SR noise.



a.

Fig. 5.3a. Probability of detection versus standard deviation τ based on y_s for the dead-zone limiter detector with sample size N = 5, signal strength A = 4, and false alarm rate $\alpha = 0.1$.



b.

Fig. 5.3b. Probability of detection versus standard deviation τ based on y_g for the dead-zone limiter detector with sample size N = 5, signal strength A = 4, and false alarm rate $\alpha = 0.1$.

5.4 Summary of Results for SR Enhanced Nonparametric Detectors

In this chapter, we have investigated the performance of the SR noise enhanced sign, Wilcoxon, and dead-zone limiter detectors. The asymptotic efficiency as well as the finite sample detection performance of the SR modified detectors are obtained. It has been shown that for the sign detector, an improvement of the asymptotic efficiency is possible under certain condition. For the Wilcox detector, there is no SR effect in terms of the asymptotic efficiency. For the dead-zone limiter detector, a similar conclusion is obtained when its c parameter is optimal. However, when c is not the optimal value, it is still possible to improve its asymptotic efficiency by adding suitable SR noise. Also, as shown in our detection example, it is possible to improve the performance of these detectors when only limited data samples are available. Specifically, a remarkable result shown for this problem is the fact that whereas the Wilcoxon detector is far superior to the sign detector without SR noise ($\tau = 0$), the presence of the optimal SR noise enables the sign detector to significantly outperform the Wilcoxon.

Overall, SR provides an important approach to enhance the performance of commonly used nonparametric detectors. Similar approaches can be employed for other nonparametric detectors. Further issues such as the determination of the optimum SR noise PDF for a larger set of nonparametric detectors are under investigation.

6.0 Application of Stochastic Resonance to Imagery

6.1 Image Quality Metrics

In Phase I, a state-of-the art assessment was conducted on image quality metrics to determine if they can provide an approximate and more efficient method of obtaining quality scores provided by human observers. Important considerations involve image noise consisting of (a) random image noise such as impulsive and additive noise as well as (b) structural noise consisting of artifacts in the imagery. Another important consideration is the degradation of image sharpness due to blurring, ringing, and blocking effects. The advantages of the image quality metrics is that they are able to dynamically monitor and adjust image quality, optimize parameters and algorithms, and benchmark image processing systems and algorithms.

Objective quality metrics can be placed in three general categories. These consist of (a) full reference, (b) no reference, and (c) reduced reference. Full reference metrics provide a so-called 'perfect version' in the sense that 'ground truth' or a 'golden image' is available as a reference. Here, the image quality is measured by comparing the difference between the reference and the distorted image. For the 'no reference' case, no such reference is available so that the image itself must be assessed. Finally, the 'partial reference' methods utilize available partial information regarding the 'perfect image' as in the case of image fusion processing where multiple images are utilized.

Several image comparison metrics have been proposed to compare the similarity between different images. Among the most prevalent is the **mean-squared error metric** which is an average of the sum of squares of the pixel value differences between each image and is expressed as

$$MSE = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}[X(i,j)-Y(i.j)]^2 \tag{6.1}$$

where $X(i,j)$ and $Y(i,j)$ are the ijth pixel values for images X and Y, respectively.

Due to its important statistical meaning, **Mutual Information (MI)** is also widely used as an evaluation measure which is given by

$$I = H(X)+H(Y)-H(X,Y) \tag{6.2}$$

where $H(\cdot)$ is the entropy. Although it can been shown that for $I = H(X)$, Y is identical to X, MI is not directly related to human perception performance; i.e., MI does not adequately indicate the visual perception performance [Chen et al. 2005]. Also, since the calculation of MI involves the estimation of a joint PDF between X and Y, MI is often not suitable for a relatively small sized image where the sample support is low. Alternatively, it is difficult to extend MI to handle a high dimensional image dataset. A related quantity, however, called the visual information fidelity metric is considered below.

Structural Similarity (SSIM) [Z. Wang, et al. 2004], a nonlinear combination of the difference in terms of the mean, variance, and contrast is also widely used and is expressed as

$$Q = \frac{1}{M}\sum_{j=1}^{M} SSIM_j \tag{6.3}$$

where

$$SSIM_j = \frac{2\overline{xy} + C_1}{(\overline{x})^2 + (\overline{y})^2 + C_1} \bullet \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}. \tag{6.4}$$

The **Visual Information Fidelity** (VIF) has also been considered [Sheikh, Bovik, et al. 2004]. It is a measure of the ratio

$$VIF = \frac{DiI}{RiI} \tag{6.5}$$

where *DiI* is the distorted image information and *RiI* is the reference image information. A signal gain and additive noise model in the wavelet domain given by

$$D = GC + V \tag{6.6}$$

is applied to describe the test image, where C denotes the random field (RF) from a wavelet sub-band in the reference image and D denotes the corresponding signal in the test image. G is a deterministic scalar gain and V is white Gaussian noise. A Gaussian distribution assumption is then made to calculate the mutual information. Overall, *DiI* and *RiI* are obtained as the summation of the **mutual information** of all sub-bands between the distorted image and the true image C and between the reference image and C, respectively. Examples of the VIF metric are shown in Fig. 6.1 for *VIF* values of 1.0 (reference image), 1.1, 0.07, and 0.10.

Reference VIF = 1.0

Contrast Enhanced VIF = 1.10

Fig. 6.1 Examples of the Visual Information Fidelity (VIF) metric (continued on next page).

Blurred VIF = 0.07 **JPEG Compressed VIF = 0.10**

Fig. 6.1 (continued) Examples of the VIF metric.

The 'No-Reference (NR) Quality Assessment' algorithms [Z. Wang, et al., 2000], [Sheikh, et al., 2005] were recently developed to evaluate image quality without a reference image; i.e., when no reference image is available. They are based on a quantization of the image distortion. Specifically, the edge sharpness level quantization relates to the edge width and the blur level estimation and noise level estimation is affected by the impulse noise level and the additive Gaussian noise level. Some existing NR metrics consist of image ringing and blurring measures, image noise level, as well as receiver operating characteristic (ROC) curves as in medical image processing where the detection performance of a certain disease is considered.

Finally, the **Adaptive Image Quality Measure** (IAQM) has been considered [Bingabr, Varshney, Farell, 2003] as a measure which provides the peak signal-to-noise ratio computed after eliminating the errors not seen by the eye and the extent (in terms of percentage of blocks) to which the image is corrupted.

In future work, we will utilize such metrics to improve the filtered image quality and determine the best parameters for the SR based approach as in cases such as the median filtered image discussed in Section 2.0. By doing so, automatic filtering algorithms will be developed and it is anticipated that performance will be improved over state-of-the-art methods.

6.2 Detection Enhancement in Imagery

We again emphasize that the specific form of the noise PDF plays a critical role to achieve enhancement via SR. We reconsider the problem addressed in [Kay 2000], but now for a two-dimensional (150x150 pixel) image. In Fig. 6.2, a 6x6 window is used to process the data so that N = 36 pixels. Fig. 6.2a shows the original image with no additive noise while Fig. 6.2b shows the image with additive Gaussian mixture noise. Fig. 6.2b is now used as the baseline observed

image containing signal and noise. In Fig. 6.2c, the sign detector processes the image of Fig. 6.2b and in Fig. 6.2d we observe a modest *image visualization* performance improvement, the result of adding white SR Gaussian noise with a variance of 2.0.

In Figs. 6.2e and 6.2f, we consider the problem from a detection viewpoint showing binary detection results for the sign detector without and with Gaussian SR noise, respectively. Here, we compare the test statistic values to a threshold to decide signal presence (white region) or signal absence (dark region). The figures demonstrate the SR effect on *detection* performance improvement. Specifically, the detection performance within the signal region has improved.

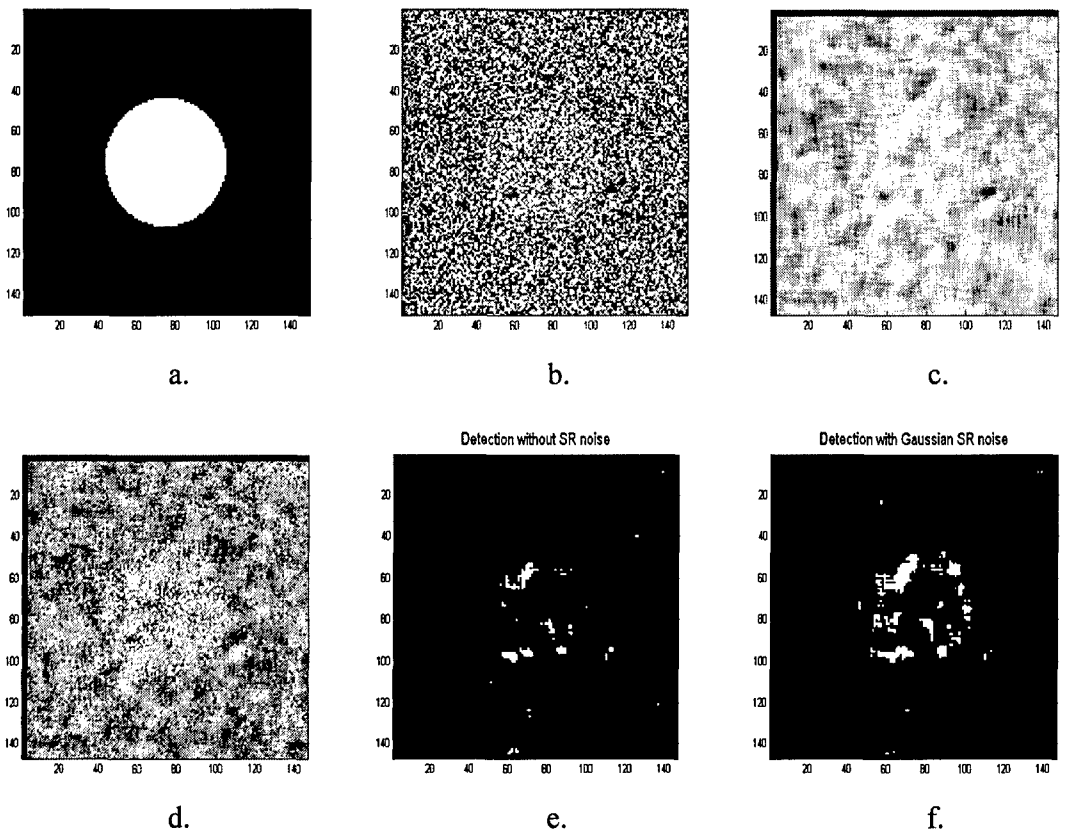


Fig. 6.2 Image enhancement using Gaussian SR white noise, a.) signal image, b.) signal image plus Gaussian mixture noise, c.) sign detector test statistic, d.) sign detector test statistic with Gaussian white SR noise, e.) detection using sign detector without SR noise, f.) detection using sign detector with Gaussian SR noise.

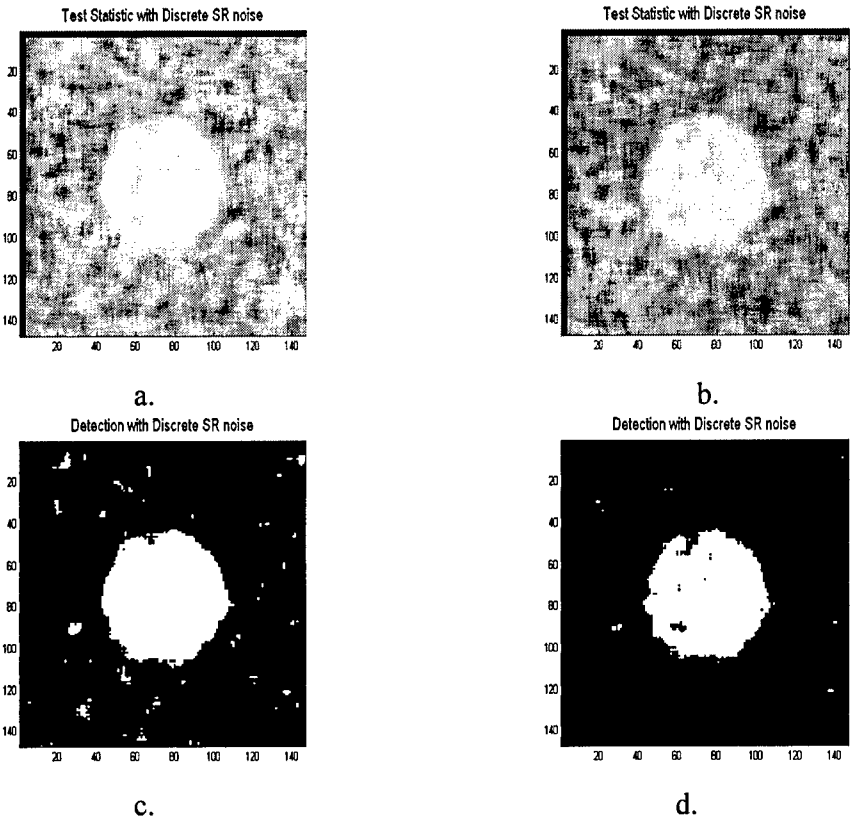


Fig. 6.3 Image enhancement using discrete SR noise PDF with $n_1 = -4.0$ and $n_2 = 2.0$, a.) test statistic values, $\lambda = 0.05$, b.) test statistic values, $\lambda = 0.1587$, c.) detection results, $\lambda = 0.05$, d.) detection results, $\lambda = 0.1587$.

In Fig. 6.3, however, we repeat the results using the discrete PDF as in (2.17), but now using $n_1 = -4.0$ and $n_2 = 2.0$. Figs. 6.3a and 6.3b show the test statistic values with probability values $\lambda = 0.05$ and 0.1587, respectively. The results reveal a noticeable improvement in the image visualization quality when compared to Fig. 6.2d. Figs. 6.3c and 6.3d show corresponding binary detection results with significant enhancement over the results for Gaussian SR noise shown in Fig. 6.2f. A comparison of Figs. 6.3c and 6.3d further reveal the impact of the parameter λ to control P_{FA}^{y} and P_{D}^{y} levels.

Finally, we consider the application of the theoretical SR detection framework to an actual image to determine the *image visualization* improvement. In Fig. 6.4, we consider the 'Lena' image with the original image shown in Fig. 6.4a and the image with a high threshold level applied in Fig. 6.4b. In practice, the latter would represent an incorrect binary threshold or perhaps a human 'eye-detector' with a damaged neuron requiring a high excitation level. In Figs. 6.4c, d, and e, specific cases of Gaussian, uniform, and optimal discrete SR noise are considered, respectively. The results demonstrate the potential for dramatic *image visualization* improvement with the application of the appropriate SR noise PDF. In phase II, attention will be given to the 'a priori' information considerations required to achieve the improvements in practice.

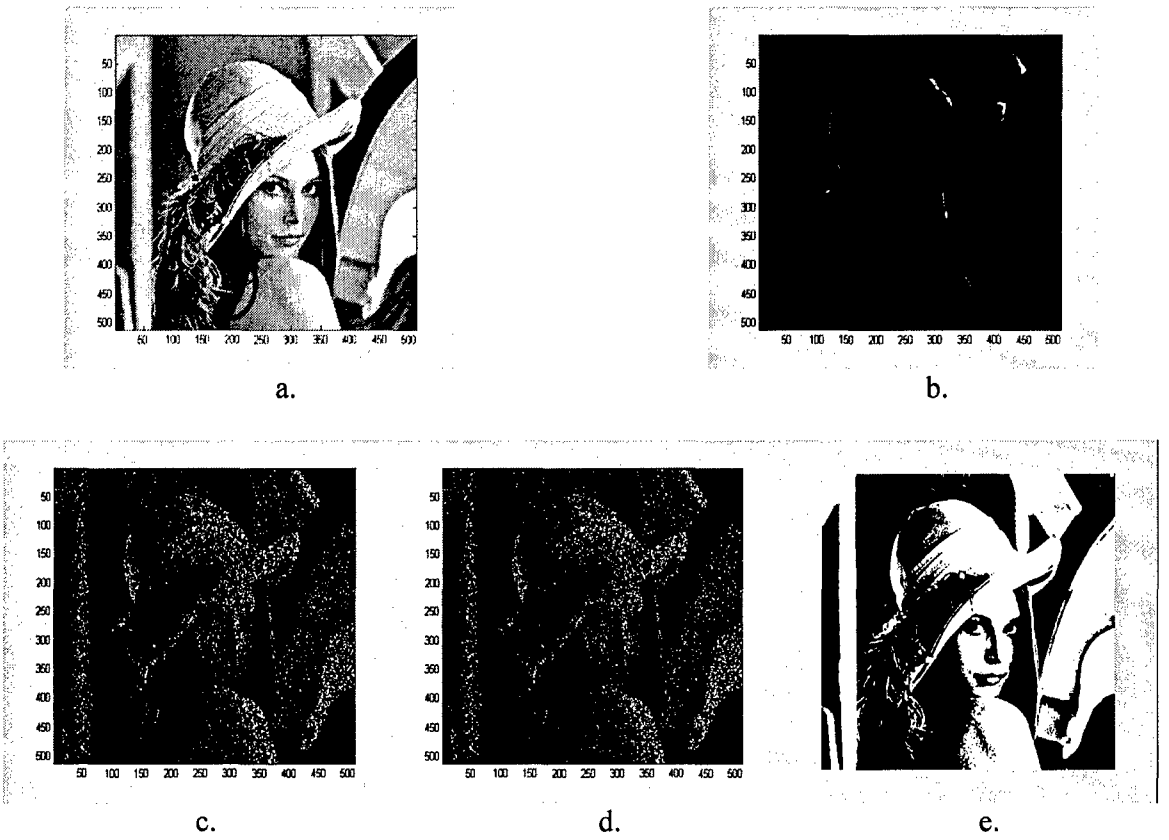


Fig. 6.4 Image visualization using the 'Lena' image a.) original Lena image, b.) image with a high threshold applied, c.) Gaussian SR noise, d.) uniform SR noise, e.) optimal discrete SR noise.

7.0 Application of Stochastic Resonance (SR) to Distributed Detection

Despite the progress achieved over the past two decades, the application of the SR effect to distributed detection has not been shown. Here, we investigate this application area for the dual hypothesis detection problem. We restrict ourselves to binary local sensor outputs, denoted by U_k, and consider the cases of both conditional independence and dependence among sensor observations. The degradation of detection performance caused by transmission errors between local sensor outputs and the fusion center is assessed. The relationship between the additive SR noise and system performance is explored. For the traditional two-stage approach using the Chair-Varshney fusion rule [Chair1986], the role of additive SR noise at both the decoding stage and the decision stage is discussed. We show that the SR phenomenon exists under certain circumstances for both cases.

7.1 Stochastic Resonance Problem Statement

We again summarize the mathematical framework here for convenience. Given a K-dimensional data vector $\mathbf{x} \in R^K$, we decide between two hypotheses H_1 and H_0,

$$H_0: p_{\mathbf{x}}(\mathbf{x};H_0) = p_0(\mathbf{x}) \tag{7.1a}$$
$$H_1: p_{\mathbf{x}}(\mathbf{x};H_1) = p_1(\mathbf{x}) \tag{7.1b}$$

where $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$ are the pdfs of $\mathbf{x}$ under H_0 and H_1, respectively. In order to make a decision, a test that can be completely characterized by a *critical function* (*decision function*) ϕ where $0 \leq \phi(\mathbf{x}) \leq 1$ for all $\mathbf{x}$, is required to choose between the two hypotheses. For any observation $\mathbf{x}$, this test chooses the hypothesis H_1 with probability $\phi(\mathbf{x})$. The detection performance of this test $\phi(\bullet)$ can be evaluated by its probability of detection P_D and probability of false alarm, P_{FA}. In order to enhance detection performance, we add SR noise to the original data process $\mathbf{x}$ and obtain a new data process $\mathbf{y}$ given by

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{7.2}$$

where the $\mathbf{n}$ is either an independent random process with pdf $p_{\mathbf{n}}(\bullet)$ or a nonrandom signal. For the case where the critical function $\phi(\bullet)$ is fixed, to improve P_D without increasing P_{FA}, the optimum SR noise has been shown to consist of no more than two discrete vectors [Chen 2006a,f]; i.e., the optimal SR noise pdf $p_{\mathbf{n}}^{opt}(\mathbf{n})$ is of the following form,

$$p_{\mathbf{n}}^{opt}(\mathbf{n}) = \lambda\delta(\mathbf{n} - \mathbf{n}_1) + (1-\lambda)\delta(\mathbf{n} - \mathbf{n}_2) \tag{7.3}$$

where $0 \leq \lambda \leq 1$, $\mathbf{n}_1$ and $\mathbf{n}_2$ are suitable K dimensional vectors.

7.2 Decision Fusion and non-ideal Transmission Channels

A typical parallel fusion model with transmission channels is shown in Fig. 7.1. For the kth local sensor, an independent binary decision u_k is made based on its observations. Without loss

of generality, assume that $u_k = 0$ if H_0 is decided and $u_k = 1$, otherwise. The detection performance of the kth local sensor node can be characterized by its corresponding probability of false alarm and detection, denoted by P_{FAk} and P_{Dk}, respectively. The kth local decision u_k is sent to the fusion center through a transmission channel C_k characterized by $p(x_k \mid u_k)$. Yhe final decision u_0 is made at the fusion center based on the received data $\mathbf{x} = [x_1, x_2, \ldots, x_K]^T$ and the fusion rule γ, i.e.,

$$U_0 = \gamma(x_1, x_2, \ldots, x_K). \tag{7.4}$$

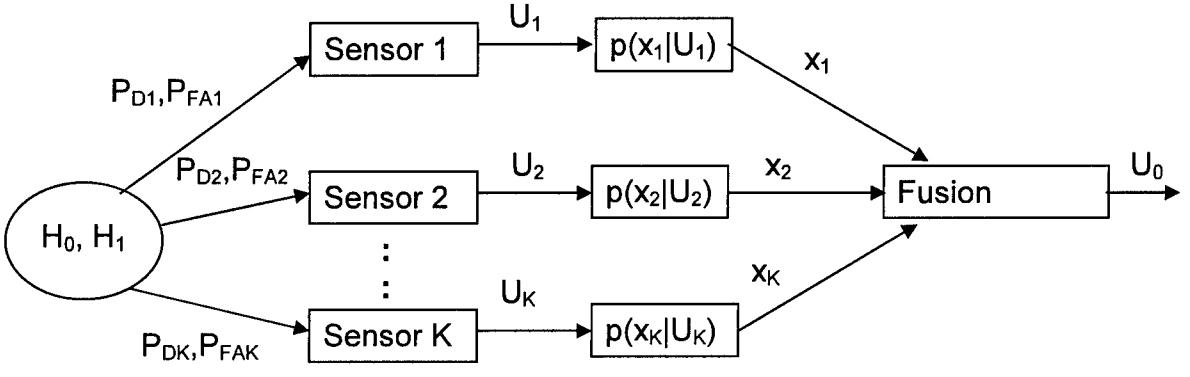


Fig. 7.1 The parallel fusion model.

In general, two different fusion rules are applicable at the fusion center depending on the different definitions of the output **x**. For the traditional two-stage approach, the output of each transmission channel x_k is the estimate of u_k; i.e., the kth channel can be described as a binary channel with crossover error probabilities α_k and β_k. The fusion rule γ_1, assuming perfect connections between the local sensors and the fusion center, is given by

$$\gamma_1 = \sum_{k=1}^{K} \left[\log \frac{(1-P_{FAk})P_{Dk}}{(1-P_{Dk})P_{FAk}} \right] x_k \underset{H_0}{\overset{H_1}{\gtrless}} \eta_1. \tag{7.5}$$

Further, by applying the channel model for the signal detection problem for the kth local sensor and its corresponding channel C_k, the relationship between x_k and the hypothesis H_i can be described as a two-layer transformation channel shown in Fig. 7.2 with its equivalent one-layer model shown in Fig. 7.3. After accounting for channel errors, the equivalent kth sensor probability of detection P_{Dk}^c and probability of false alarm P_{FAk}^c are given by

$$P_{Dk}^c = (1 - \beta_k)\, P_{Dk} + \alpha_k(1 - P_{Dk}) \tag{7.6}$$

and

$$P_{FAk}^c = (1 - \beta_k)\, P_{FAk} + \alpha_k(1 - P_{FAk}). \tag{7.7}$$

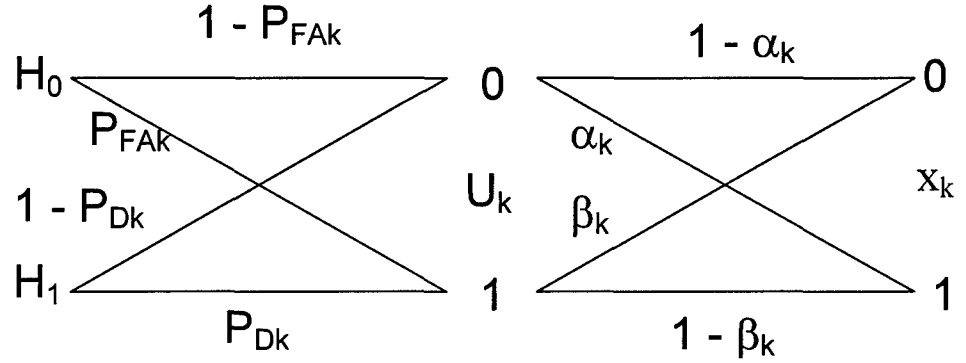


Fig. 7.2 A two layer transmission channel model for a distributed detection system.

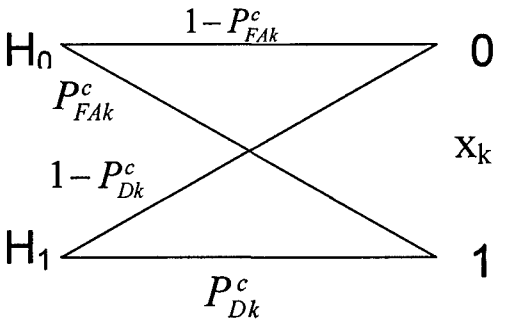


Fig. 7.3 Channel model for a signal detection problem in local sensor k.

Therefore, when the channel statistics α_k and β_k are available, the optimum fusion rule γ_2 at the fusion center is given by the Chair-Varshney rule [Chair 1986]

$$\gamma_2 = \sum_{k=1}^{K}\left[\log\frac{(1-P^c_{FAk})P^c_{Dk}}{(1-P^c_{Dk})P^c_{FAk}}\right]x_k \underset{H_0}{\overset{H_1}{\gtrless}} \eta_2 . \tag{7.8}$$

When the channel is perfect, i.e., $\alpha_k = \beta_k = 0$, we have $P^c_{Dk} = P_{Dk}$ and $P^c_{FAk} = P_{FAk}$ while $\gamma_2 \equiv \gamma_1$.

Compared to the two-stage fusion approach where the output of channel k is the binary estimate of u_k, channel aware decision fusion rules have been developed recently [B. Chen 2004], [Niu 2006] based on the direct observation of channel data. In this approach, the output of channel C_k is no longer a binary variable but a continuous random variable. The exact form of $p(x_k \mid u_k)$ depends on the coding rule at each local sensor and its corresponding noisy channel model. For example, for many wireless sensor network scenarios, the kth channel between the kth local sensor and the fusion center can be modeled as a unit power Rayleigh fading channel consisting of additive Gaussian noise with variance σ_k^2. Assuming that the local sensors send a 1 when H_1 is decided and -1 otherwise, the channel statistic $p(x_k \mid u_k)$ is given by

$$p(x_k \mid u_k = 1) = \frac{2\sigma_k}{\sqrt{2\pi}(1+2\sigma_k^2)} \exp(-\frac{x_k^2}{2\sigma_k^2}) \left[1 + \sqrt{2\pi} a_k x_k \exp(a_k^2 x_k^2 / 2) Q(-a_k x_k)\right], \quad (7.9)$$

and

$$p(x_k \mid u_k = 0) = \frac{2\sigma_k}{\sqrt{2\pi}(1+2\sigma_k^2)} \exp(-\frac{x_k^2}{2\sigma_k^2}) \left[1 - \sqrt{2\pi} a_k x_k \exp(a_k^2 x_k^2 / 2) Q(a_k x_k)\right], \quad (7.10)$$

where $a_k = \frac{1}{\sigma_k \sqrt{(1+2\sigma_k^2)}}$ and $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt$ is the complimentary distribution function of the standard Gaussian distribution.

Several decision fusion rules that require different degrees of *a priori* knowledge have been proposed in [B. Chen 2004], [Niu 2006]. We summarize the test statistics for a few of them here.

1. Chair-Varshney Fusion Rule

$$\gamma_3 = \sum_{k=1}^{K} \left[\log \frac{(1-P_{FAk})P_{Dk}}{(1-P_{Dk})P_{FAk}} \right] I(x_k) \quad (7.11)$$

where $I(\cdot)$ is an indicator function

$$I(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

2. Equal Gain Combining (EGC) Fusion Statistic

$$\gamma_4 = \frac{1}{K} \sum_{k=1}^{K} x_k, \quad (7.12)$$

3. Likelihood Ratio Test Based on Channel Statistics (LRT-CS)

These test statistics are based on the knowledge of channel statistics and local detection performance indices

$$\gamma_5 = \sum_{k=1}^{K} \log \left\{ \frac{1 + \sqrt{2\pi} a_k x_k \exp(a_k^2 x_k^2 / 2) Q(-a_k x_k)}{1 - \sqrt{2\pi} a_k x_k \exp(a_k^2 x_k^2 / 2) Q(a_k x_k)} \right\} \quad (7.13)$$

It has been shown that although γ_3 is near optimal when the channel SNR is high, it suffers significant performance loss at low to moderate channel SNR. However, as shown in the next section, the detection performance of γ_3 can be improved by adding independent SR noise.

7.3 Noise Enhanced Decision Fusion

We now consider two examples to demonstrate the possible SR effect is decision fusion. Consider the first decision fusion approach where an estimate of u_k is obtained before being sent to the fusion center. Here, two sensors are involved in the system. For sensor 1, we assume that $P_{D1} = 0.8$, $P_{FA1} = 0.1$ and the detection performance for sensor 2 is $P_{D2} = 0.95$, $P_{FA2} = 0.05$. We further assume that channel one is a perfect channel while channel 2 is a noisy channel with crossover error probabilities $\alpha_2 = \beta_2 = 1/3$. In practice, this may depict a scenario in which sensor I is located far from the signal source, but close to the fusion sensor. On the other hand, sensor 2 may be located close to the signal source, but far from the fusion center. Therefore, at the fusion center, the detection performance of the second sensor is actually equivalent to $P_{D2}^c = 0.65$ and $P_{FA2}^c = 0.35$. The detection performance of fusion rules γ_1 and γ_2 is shown in Fig. 7.4. Clearly, due to the performance loss in the noisy channel, γ_1 is no longer the optimum fusion rule and its detection performance is degraded. In order to improve the performance of γ_1, we add SR noise to the observed data x_2 to obtain a new data sample y_2. Since x_2 is a discrete random variable, we use the noisy binary channel model with crossover probabilities α_{SR} and β_{SR} to generate the new noisy SR data sample y_2. Thus, the procedure here is to utilize the cross-over error information in the enhancement procedure. Specifically, x_2 is observed at the fusion center. The decision x_2 is either retained or changed depending upon the outcome of a comparison of a uniformly distributed random variable w with some specified values α_{SR} and β_{SR}. If x_2 is observed as a zero and $w \leq \alpha_{SR}$, then x_2 is changed to a one. Conversely, if x_2 is observed as a one and $w \leq \beta_{SR}$, then x_2 is changed to a zero. It is interesting to note that the randomized procedure actually introduces additional errors in sensor 2, but places more emphasis on sensor 1 with the negligible channel cross-over errors.

The fusion performance of γ_1 using the new data sample y_2 is also plotted in Fig. 7.4. When $\alpha_{SR} = 0$ and $\beta_{SR} = 0.5$, a higher P_D for this SR modified fusion system is observed for $P_{FA} \in [0.07, 0.35]$ when compared to the original γ_1. A similar effect is also observed for the parameter setting with $\alpha_{SR} = 0.5$ and $\beta_{SR} = 0$. Furthermore, it can be shown that performance enhancement for the shaded region in Fig. 7.4 is possible by adding suitable SR noise.

In the next example, in order to examine the possible SR effect in decision fusion for a wireless sensor network, we choose the number of sensors to be K = 8, with $P_{Dk} = 0.5$ and $P_{FAk} = 0.05$ for each sensor. The SR noise n is chosen to be a DC value of A; i.e., instead of using the original data x_k to perform decision fusion using γ_3, new SR modified data $y_k = x_k + A$ is used. Due to the computational complexity of this detection problem, the detection performance evaluation is obtained by intensive Monte Carlo evaluation. Fig. 7.5 shows the deflection measures [Picinbono 1995] for different fusion rules with SR noise A = -0.2 and the deflection measure is defined as

$$D(\gamma) = \frac{\left[E(\gamma \mid H_0) - E(\gamma \mid H_1)\right]^2}{Var(\gamma \mid H_0)}. \tag{7.14}$$

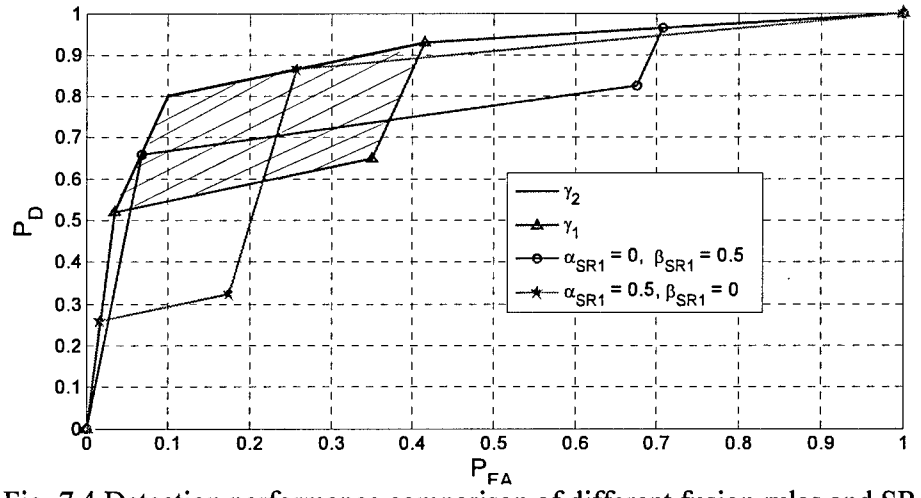


Fig. 7.4 Detection performance comparison of different fusion rules and SR noise.

As observed in Fig. 7.5, for most values of channel SNR, by adding a stochastic resonance noise n = A = -0.2 to the observed data x, the deflection coefficient is improved. An interesting observation is that when channel SNR is between 10dB and 20dB, the deflection coefficient of SR modified γ_3 is even higher than that of LRT-CS. However, this does not imply that the detection performance of SR modified γ_3 is better than LRT-CS since these test statistics are not Gaussian distributed.

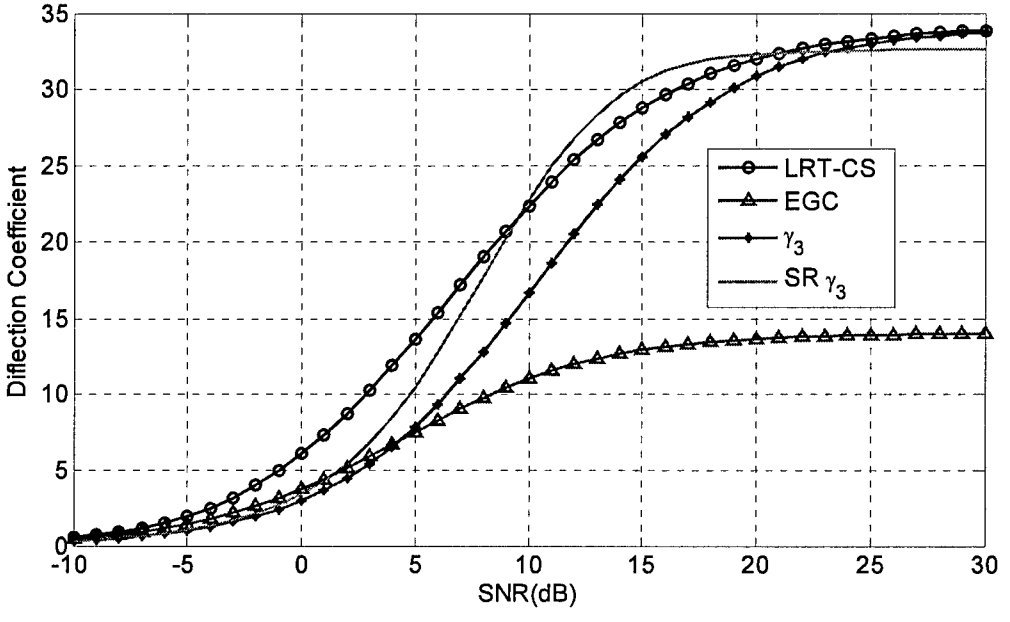


Fig. 7.5 Deflection Coefficient for different test statistics, A = -0.2.

For a fixed channel SNR and SR enhanced γ_3 decision fusion, the relationship between different values A and deflection coefficient D is shown in Fig. 7.6. As A starts becoming negative, the deflection coefficient D first increases and then, after attaining its peak, decreases as A decreases further. The optimum value of A, namely A_0, which maximizes D is an increasing function of SNR. When SNR is very high, we have $A_0 \approx 0$ which is consistent with the conclusion drawn in [B. Chen 2004], [Niu 2006] where the asymptotic optimum of γ_3 is derived; i.e., SR noise will not improve performance at high SNR.

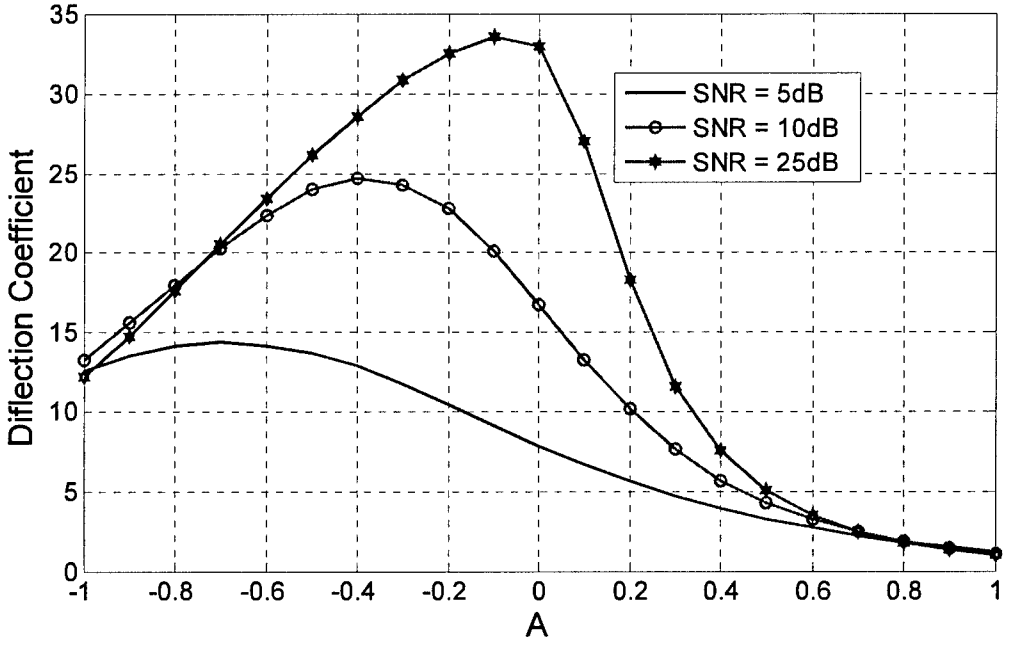


Fig. 7.6 Deflection coefficient for SR enhanced γ_3 decision fusion for different channel SNR.

Fig. 7.7 gives the ROC curves corresponding to different fusion statistics at channel SNR of 5dB. Clearly, the SR modified γ_3 fusion rule provides a better detection performance than both the EGC and the original γ_3 rule.

To explain this SR enhanced detection phenomenon, we first obtain the relationship between the Rayleigh fading channel model and the binary channel model. Corresponding to the transmission channel illustrated in Fig. 7.2, we have for γ_3,

$$\alpha_k = p(I(y_k) = 1 | u_k = 0), \tag{7.15}$$

and

$$\beta_k = p(I(y_k) = 0 | u_k = 1) = 1 - p(I(y_k) = 1 | u_k = 1). \tag{7.16}$$

From (9) and (10), after some calculation, we have

$$p(I(y_k) = 1 | u_k = 1) = p((x_k + A) \geq 0 | u_k = 1)$$

$$= p(x_k \geq -A | u_k = 1)$$
$$= \int_{-A}^{\infty} \frac{2\sigma_k}{\sqrt{2\pi}(1+2\sigma_k^2)} \exp(-\frac{t^2}{2\sigma_k^2}) \left[1 + \sqrt{2\pi} a_k t \exp(a_k^2 t^2/2) Q(-a_k t)\right] dt$$
$$= Q\left(-\frac{A}{\sigma_k}\right) + \frac{Q(Aa_k)}{\sqrt{1+2\sigma_k^2}} \exp\left(-\frac{A^2}{1+2\sigma_k^2}\right) \quad (7.17)$$

and

$$p(I(y_k) = 1 | u_k = 0) = 1 - Q\left(\frac{A}{\sigma_k}\right) - \frac{Q(-Aa_k)}{\sqrt{1+2\sigma_k^2}} \exp\left(-\frac{A^2}{1+2\sigma_k^2}\right). \quad (7.18)$$

From (17) and (18), it can be shown that α_k monotonically decreases and β_k monotonically increases as A decreases. An illustration of such relationship is shown in Fig. 7.8 for the case of channel SNR = 5dB. Also, from (6) and (7), it can be shown that for any fixed channel SNR and probability of false alarm P_{FA}, the probability of detection P_D given by the SR modified fusion rule γ_3 is determined by the crossover error probabilities α_k and β_k, k = 1, 2, ..., K which are functions of A. Therefore, there exists a suitable A which yields the best detection performance, i.e., maximizes the P_D for a given P_{FA}. When A = 0, $\alpha_k = \beta_k = \frac{1}{2} - \frac{1}{2\sqrt{1+2\sigma_k^2}}$. When SNR is very high, α_k, $\beta_k \rightarrow 0$, and the channel C_k becomes a near perfect channel. As a result, γ_3 becomes a near optimum fusion rule.

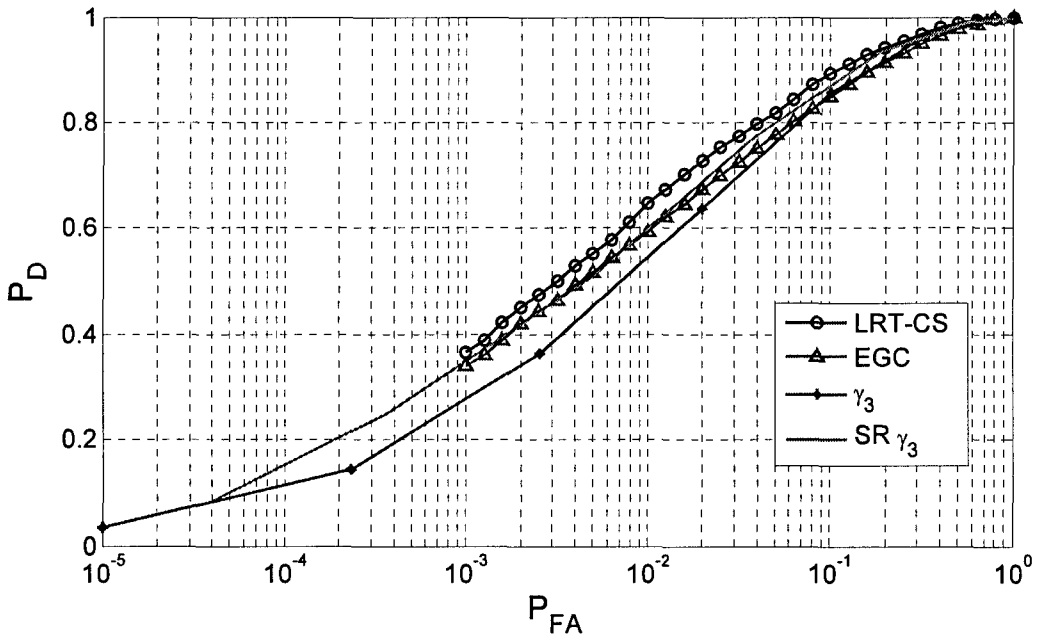


Fig. 7.7 ROC curves for various fusion statistics; SNR = 5dB, A = -0.2.

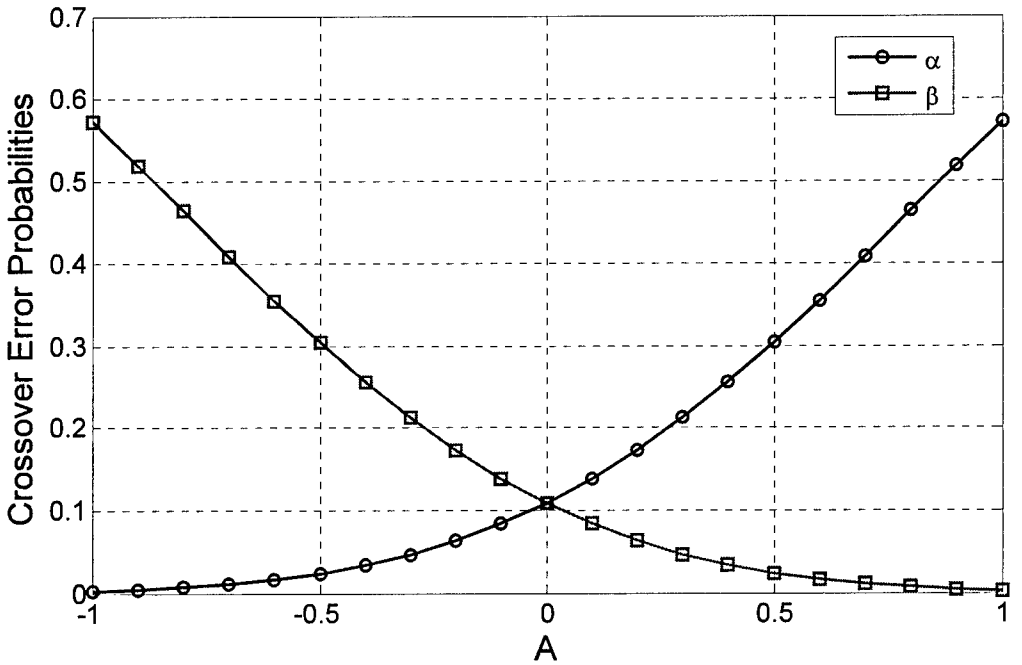


Fig. 7.8 The equivalent channel crossover errors as a function of A, SNR = 5dB.

7.4 Summary of Results for Distributed Detection using SR

In this section, we have considered the detection performance of distribute detection and fusion systems in the presence of non-ideal transmission channel. For fusion of decisions transmitted over channels that can be modeled as a binary channel, we showed that the detection performance of some decision fusion systems can be improved by randomly changing the received binary signal, i.e., by adding stochastic resonance noise. For the problem of fusion of decisions transmitted through a Rayleigh fading channel, we established the equivalence between this fading channel and the binary channel model for the widely used two-stage Chair-Varshney fusion rule. We further demonstrated the existence of the SR phenomenon in this fusion problem by adding a discrete DC value to the observed signal at the fusion center. A significant improvement of detection performance is reported when suitable noise is selected. Further results including an adaptive approach to learn the optimum noise vale n and an extension of this SR effect to other decision fusion systems will be forthcoming.

8.0 Optimal Decision Processing by the Transformation Method

For decision problems whose decision region is suboptimal, we now show how to transform the decision statistic to recover optimal performance [Kay, 2006a]. This novel methodology holds considerable promise in a wide range of application areas involving decision theory. The procedure simply amounts to transforming the decision statistic to yield a combined statistic/decision region which is optimal. The approach may be thought of as a generalization of the stochastic resonance phenomenon, which employs a random linear transformation, and hence should be widely applicable to many practical problems. Maximization of the probability of a correct decision, as considered here for example, has direct applications to communication theory. The method considered here treats the univariate case. Its generalization to the multivariate case will be considered in future work.

8.1 Mathematical Description of the Transformation Method

A mathematical justification of the approach is given in this subsection with an example in the next subsection. Here, we assume that the problem is to decide between two hypotheses H_0 and H_1 based on the observed scalar test statistic x. The two hypotheses are assumed to be random events with prior probabilities of π_0 and π_1. This test statistic is a function of the original data. A future problem will address the extension to the case when the original data is accessible. Based on the observed data sample x, a decision rule has been implemented as follows:

$$\phi(x) = 1 \quad \text{decide } \mathcal{H}_1 \tag{8.1a}$$
$$\phi(x) = 0 \quad \text{decide } \mathcal{H}_0. \tag{8.1b}$$

This decision rule is assumed to be *suboptimal.* Denoting the probability density functions (PDFs) as $p_0^X(x)$ and $p_1^X(x)$ under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, the probability of a correct decision is

$$\begin{aligned} P_c &= \pi_0 \int_{-\infty}^{\infty} (1-\phi(x)) p_0^X(x)dx + \pi_1 \int_{-\infty}^{\infty} \phi(x) p_1^X(x)dx \\ &= \pi_0 + \int_{-\infty}^{\infty} \phi(x)\left[\pi_1 p_1^X(x) - \pi_0 p_0^X(x)\right]dx \end{aligned} \tag{8.2}$$

Note that unless $\phi(x) = 1$ for all x such that that $\pi_1 p_1^X(x) - \pi_0 p_0^X(x) > 0$ and zero otherwise (which is the optimal decision rule) this probability will not be maximized.

Now consider that we transform the test statistic as $y = g(x)$ using some function g. The function is assumed to be piece-wise monotonic so that over each interval of a finite number of disjoint intervals either $g'(x) \geq 0$ or $g'(x) < 0$ (the prime denotes differentiation). The transformed test statistic y is then input to the decision rule $\phi(\bullet)$ to yield $\phi(y)$. This is in accordance with the assumption that the decision rule is fixed and so cannot be changed. Only the test statistic can be modified. We will see, however, that this is mathematically equivalent to modifying the decision rule. To do so note that P_c, which is now based on y, is

$$P_c = \pi_0 + \int_{-\infty}^{\infty} \phi(y)\left[\pi_1 p_1^Y(y) - \pi_0 p_0^Y(y)\right] dy. \tag{8.3}$$

We now utilize the piece-wise monotonic assumption of g(x) to write the real line as $R = \cup_{i=1}^{N} I_i = \cup_{i=1}^{N} (a_i, b_i)$, where $a_1 < b_1 < a_2 < b_2 ... < a_N < b_N$ and the intervals are open. The function g(•) is monotonic over each interval I_i. The omission of a finite number of points from the integral will not affect the results as long as the PDFs do not contain any impulses at these points (or cumulative distribution function is continuous over all of R). Now we have that by defining

$$J_i = \begin{cases} (a_i, b_i) & \text{if } g'(I_i) \geq 0 \\ (b_i, a_i) & \text{if } g'(I_i) < 0 \end{cases} \tag{8.4}$$

and using a change of variables from y to g(x)

$$\begin{aligned} P_c &= \pi_0 + \int \phi(g(x))\left[\pi_1 p_1^Y(g(x)) - \pi_0 p_0^Y(g(x))\right] g'(x) dx \\ &= \pi_0 + \sum_{i=1}^{N} \int_{J_i} \phi(g(x))\left[\pi_1 p_1^Y(g(x)) - \pi_0 p_0^Y(g(x))\right] g'(x) dx. \end{aligned} \tag{8.5}$$

Note that for the intervals for which $g'(x) < 0$, we have $J_i = (b_i, a_i)$. Absorbing the negative sign into $g'(x)$ for the monotonically decreasing function intervals yields

$$P_c = \pi_0 + \sum_{i=1}^{N} \int_{J_i} \phi(g(x))\left[\pi_1 p_1^Y(g(x)) - \pi_0 p_0^Y(g(x))\right] |\, g'(x) \,| \, dx. \tag{8.6}$$

Next we recognize that $p_1^Y(g(x)) \,|\, g'(x) \,| = p_1^X(x)$ and $p_0^Y(g(x)) \,|\, g'(x) \,| = p_0^X(x)$, so that we finally have

$$\begin{aligned} P_c &= \pi_0 + \sum_{i=1}^{N} \int_{J_i} \phi(g(x))\left[\pi_1 p_1^X(x) - \pi_0 p_0^X(x)\right] dx \\ &= \pi_0 + \int_{-\infty}^{\infty} \phi(g(x))\left[\pi_1 p_1^X(x) - \pi_0 p_0^X(x)\right] dx. \end{aligned} \tag{8.7}$$

We now have that the probability of a correct decision is based on ϕ(g(x)). In effect by *transforming the decision test statistic x to g(x) we have been able to effectively modify the decision region.* It is clear from (8.7) that for optimal performance we must have

$$\phi^*(x) = \phi(g(x)) = \begin{cases} 1 & \text{if } \pi_1 p_1^X(x) > \pi_0 p_0^X(x) \\ 0 & \textit{otherwise} \end{cases} \tag{8.8}$$

where $\phi^*(x)$ denotes the optimal decision rule based on x. In composite function notation, we require for optimality that

$$\phi \circ g(x) = \phi(g(x)) = \phi^*(x). \tag{8.9}$$

We need only determine the function g(•). We provide an example in the next section.

8.2 An Example of the Transformation Method

We now consider a very simple example, for which the solution is obvious. The example is that of deciding between $x \sim \mathcal{N}(0,1)$ under $\mathcal{H}_0$ and $x \sim \mathcal{N}(1,1)$ under $\mathcal{H}_1$, where $\mathcal{N}(\mu,\sigma^2)$ denotes a Gaussian PDF with mean μ and variance σ^2. The prior probabilities are $\pi_0 = \pi_1 = ½$. We assume that the *suboptimal* decision rule is to decide $\mathcal{H}_1$ if $x \geq 0$ and decide $\mathcal{H}_0$ if $x < 0$. The optimal decision rule for this problem is to decide $\mathcal{H}_1$ if $x \geq 1/2$ and decide $\mathcal{H}_0$ if $x < ½$. This is just the maximum likelihood decision rule [Kay 1998]. The suboptimal decision rule produces the correct decision for all x not in the interval [0, 1/2). It is clear now that to modify the suboptimal decision rule to make it optimal we need only map the values of x in the interval [0,1/2) into any other interval for which the suboptimal decision rule will produce a zero at its output. For example, we could use

$$g(x) = \begin{cases} x & \text{for } x \geq 1/2 \text{ and } x < 0 \\ -x & \text{for } 0 \leq x < 1/2 \end{cases}. \tag{8.10}$$

Note that the effect of the transformation is to do nothing (g(x) = x) if the test statistic value will produce the correct decision. However, in the interval [0,1/2) the decision is incorrect. To convert it to a correct decision, we simply negate the value of the test statistic as g(x) = -x. Then, the values $0 \leq x \leq ½$ become negative and are decided to correspond to $\mathcal{H}_0$ in accordance with the suboptimal decision rule. Finally, it should be observed that the function chosen is piece-wise monotonic (as well as discontinuous).

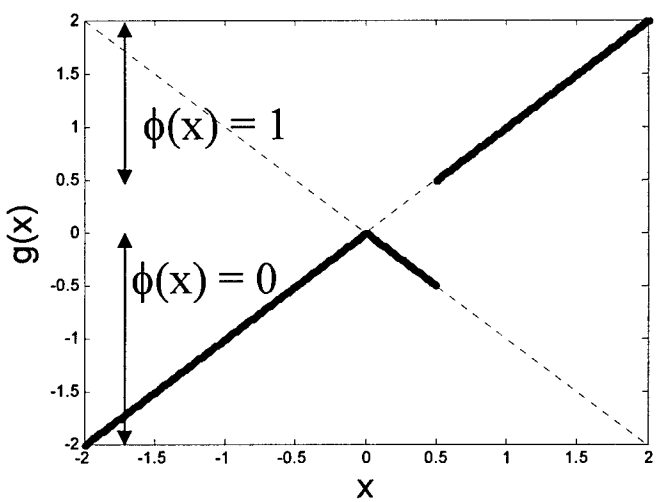


Figure 8.1 Transforming Function – One of many possibilities.

9.0 Recommendations and Future Considerations

9.1 Extensions to the Optimized SR Detection and Estimation Framework

a. Optimized SR Detection Framework for the Unknown PDF Problem

In our Phase I effort, as discussed earlier, we have addressed the problem of using SR to improve detection performance for the case when the probability density functions (PDFs) under the two hypotheses are known. In future work, the emphasis must now be placed on the problem of adaptive learning and training of SR modified detectors in the unknown PDF case. For any specific test, we assume that the prior probabilities of both hypotheses are known and that we have the knowledge of the model of this detection problem. Therefore, given a set of testing samples, the underlying PDFs for the two hypotheses detection problem may be determined by statistical learning methods [Vapnik 1996]. Furthermore, this learned PDF/CDF can be employed further to construct a better detector or to improve the detection performance of the detector obtained by the SR approach introduced in two of our papers [Chen, et. al., 2006a, b] developed during Phase I. Other model based learning algorithms such as the Ozturk algorithm [Ozturk, 1992] developed at Syracuse University (used to learn the distribution model) and the EM algorithm (used to learn the parameters of a certain model) shall also be explored. The information extracted from data via these methods will be further utilized to determine the form of the optimum SR noise distribution for the detection problem under consideration.

b. Variable Decision Functions

In our prior work in Phase I, the test statistics as well as the detection threshold were assumed fixed. Extension of the SR formulation of the fixed test statistic and detection threshold case to the variable threshold and variable statistic case will be pursued. Let Φ denote the set of all possible test statistics, ϑ be the set of all possible thresholds and Π be the set of all possible SR noise PDFs. For a Neyman-Pearson detection approach, we know that for any fixed probability of false alarm P_{FA}, its corresponding P_D can not be greater than 1. Thus, there must exist at least one $(\phi,\theta,\pi)\in(\Phi,\vartheta,\Pi)$ that maximizes the P_D with false alarm rate less than or equal to P_{FA}. The same conclusion can also be drawn for the Bayesian case. From our work reported in [Chen, et al., 2006a] and [Kay, et al., 2006], we have: $\pi = p_n(x) = \delta(x-c)$ for the Bayesian approach and $\pi = p_n(x) = \lambda\delta(x-d) + (1-\lambda)\delta(x-f)$ for the NP approach. The next step of our proposed Phase II effort in this area is to establish some relationships among ϕ, θ and π for the variable decision function and to further simplify the procedure to find the best SR detection systems.

c. Optimum SR Noise Considerations under Constraints

In some practical applications, additional constraints on SR noise may be applicable. For example, in a wireless sensor network (WSN), the power consumption at each sensor is a concern. In this case, we would want to know the form of optimum SR noise PDF if the detector

or sensor's power is limited. This important case of SR noise under potential constraints will be investigated.

d. Optimum SR Noise for the Multiple Hypotheses Test

In our prior work in Phase I, we focused our attention on binary hypothesis testing problems with considerable success. We now intend to extend our results to the multiple hypotheses testing problem. This problem is encountered in classification problems in various applications such as target identification, medical imaging, and remote sensing. As is well known from statistical decision theory, due to the complexity of the problem and unlike the two hypotheses test, the optimum solution of a multiple hypotheses test is very difficult to determine. The form of the optimum decision regions is not rectangular in general and in many cases non-optimum rectangular decision regions are employed. Therefore, our SR approach may play a very important role in this case. The role of SR has not been investigated in this context. In future analyses, we will seek to determine the form of the optimum SR noise distributions for this important case.

e. Robust Nonlinear Systems and Robust SR Noise

In the Phase I study, we evaluated the detection performance of certain detectors by adding independent SR noise provided the PDF for each hypothesis is known and fixed. In the next step of our research, a set of nonlinear systems will be considered and their robustness subject to small variations of the input signals will be evaluated. The relationship between the nonlinear system and SR noise from the viewpoint of robustness shall be explored.

f. SR Enhanced Signal Estimation

Dithering related techniques (in the SR sense) have been widely used in preserving signal information before AC to DC quantization or some other forms of transformations (preprocessing) [Wannamaker, 2000]. The quantized/transformed signal is then used as the input to various signal processing systems (post processing). However, due to the fact that many signal processing systems are nonlinear, the quantized/transformed signal may not be the optimal input for certain kinds of systems. This is mainly due to the following two reasons: 1. The SR noise in the preprocessing part may not be optimum. 2. There may exist a mismatch between the SR noise modified preprocessed signal and the signal processing system that follows; i.e., the SR noise and its related preprocessed signal may be the optimum signal for one post processing system, but may not be optimum for another. Thus, it is a very important and interesting problem as to how to tune the preprocessed signal again by adding SR noise or by some other methods. As an example, in some preliminary work, we have demonstrated the existence of such an SR phenomenon in a system where maximum likelihood estimation (MLE) is applied to estimate the signal parameter from the 1-bit quantized data. This notion will be investigated in more detail for a wider class of problems and a systematic theory for estimation problems will be developed just as we have initiated the theoretical formulation for detection problems in our Phase I work.

g. Numerical Approaches to Optimum SR Noise PDF Determination

Numerical optimization approaches will be pursued to obtain an optimal solution of SR noise PDF. As shown in our previous results, the theoretical optimum SR noise PDF is shown to be a single impulse for the Bayesian approach and two-peak noise for the NP approach subject to a constant false alarm rate (CFAR) constraint. It is relatively easy to obtain the solution for the single dimensional problem. However, for higher dimensional problems, it is much more difficult to find the solution. Thus, a set of numerical approaches including but not limited to Genetic Algorithm, Simulated Annealing and Particle Swarm Optimization techniques need to be examined for potential use for this problem. We plan to conduct a study and characterize the efficacy of different optimization algorithms for SR problems.

h. Nonlinear Stochastic Resonance

The concept of SR will be extended to a much broader context. So far, we have restricted ourselves to the additive form of SR noise; i.e., we have only considered $\mathbf{y} = \mathbf{x} + \mathbf{n}$. Here, we will consider the case where the noise can actually take any functional form, i.e., we will extend it to the more complicated case where y is a function of input x and noise n, such that $y = f(x,n)$. For example, we may consider a SR noise modified observation model to be multiplicative, i.e., $y = n*x$. Further, instead of considering only one source of noise, multiple noise sources can be considered. For example, we could have a multiplicative part and the other an additive part. One such kind of noise model is applicable when a signal is transmitted over a Rayleigh fading channel where $y = n_1*x + n_2$. The detection performance, the optimum solution and the performance bounds for these more general SR schemes will be investigated with improved performance anticipated.

i. Detection Enhancement in the Presence of Correlated Non-Gaussian Noise

Further consideration should also be given to performance improvement to be realized by SR applied to nonparametric detectors for the problem of signal detection in correlated non-Gaussian noise and additive white Gaussian noise. Here, the bimodal models considered in Phase I shall be generalized to the multi-modal compound Gaussian model. Further, a very general class of non-Gaussian processes known as Spherically Invariant Random Processes (SIRP) [Yao, 1973] shall be considered. The theory of SIRPs provides an elegant and mathematically tractable approach for the generation of multivariate non-Gaussian PDFs. Issues such as detection performance robustness and estimation efficiency (i.e., sample support size) are essential for these analyses.

There are two types of models for correlated non-Gaussian processes: (1) the endogenous model and (2) the exogenous product model [Conte, 1987]. For the endogenous model, the desired non-Gaussian PDF and correlation function is realized using a zero memory, non-linear transformation on a real Gaussian process. In this approach, however, it is not possible to control both the PDF and the correlation independently [Rangaswamy, 1993]. Further, the nonlinear transformation may give rise to non-Gaussian processes with a non-negative definite covariance matrix. For the exogenous model, however, the desired non-Gaussian process is generated by the product of a Gaussian random process and an independent non-Gaussian process which can be correlated. Thus, the PDF and correlation can be independently controlled.

This feature provides an important capability for both detection and estimation performance assessments. Issues such as detection performance robustness and estimation efficiency (i.e., sample support size) are of prime interest for these analyses.

9.2 Optimal Decision Processing by the Transformation Method

The derivation of the transformation method discussed in Section 4.0 will be extended to the multivariate case in the future. Specifically, the more general case of data vectors of length N shall be considered. This will allow the application of the approach to the more general problem when multi-sample data is available. The choice of the g(•) function is the critical issue here. In theory, it is always possible to choose an appropriate function to implement the optimal mapping. However, in practice, if we are within a high-dimensional space, it may not be obvious how it is chosen. We plan to investigate this approach as well as a possible extension to the case of random transformations. The usual SR method may be thought of as a random transformation that is effected by adding a random variable to the data. However, the restriction to an addition means that the transformation is constrained to be linear. Clearly, this is so restrictive as to impede the possible implementation of an optimal decision rule. Our approach will alleviate this restriction.

9.3 Stochastic Resonance in Imagery

a. Image Quality Metrics

As noted previously, important image processing considerations involve image noise consisting of (a) random image noise such as impulsive and additive noise as well as (b) structural noise consisting of artifacts in the imagery. Another important consideration is the degradation of image sharpness due to blurring, ringing, and blocking effects. The advantages of the image quality metrics is that they are able to dynamically monitor and adjust image quality, optimize parameters and algorithms, and benchmark image processing systems and algorithms.

Future work should further consider the image metrics noted in Section 6.1. We shall utilize the appropriate metrics to assess improvement of the filtered image quality and determine the best parameters for the SR based approaches as in cases such as the median filtered image discussed in Section 7.0. By doing so, automatic filtering algorithms will be developed and it is anticipated that performance will be improved over current state-of-the-art methods.

b. Noise Reduction

In our prior work in Phase I, it has been shown that SR can help to improve the noise filtering performance in image processing. In particular, in [Chen, et. al., 2006d], we have shown that SR noise can improve the performance of median filtering. This provides us the motivation to apply SR to a much broader class of image filtering and restoration problems for improved human visualization. In unreported work, we have further demonstrated the existence of the SR effect for an image corrupted by Gaussian noise. Our preliminary experiments show that the performance of Wiener filters improved when suitable SR noise is added. This is very encouraging since Wiener filters are optimal for the Gaussian noise case with a stationary signal. The reason for such improvement is the fact that the assumption of stationarity does not usually

hold in images even for a small local region and almost all images are non-stationary, i.e., their content varies depending on location. In the future, we should focus on the derivation of the optimum solution of the SR noise PDF for a large class of image enhancement problems. Emphasis should be placed on the development of efficient algorithms, possibly including an adaptive self-learning algorithm, to find the optimum SR noise for image filtering applications.

c. Image fusion

Image fusion, an important image processing technique to integrate useful information from different input images, has been widely used in a number of applications such as remote sensing and medical image processing. Among all of the image fusion algorithms, ***wavelet based multi-resolution image fusion*** is the most popular and powerful approach. One critical and difficult aspect in such an approach is to determine the fused image wavelet coefficient values. A number of algorithms have been developed to tackle this problem. Future work should focus on how to use SR to improve image fusion quality, namely, how the SR noise can play a role to better select the wavelet coefficients as well as to determine the suitable neighborhood for an improved fused image. The image quality will be evaluated based on image quality metrics (see Section 6.1) that we have developed based on the human visualization system. Thus, a SR based image fusion framework will be developed and its performance evaluation will be carried out.

d. Target Detection and Identification in Images

An important future consideration for SR enhanced detection is the extension of the current techniques from single pixel detection (the circular image of Fig. 6a) to more general target detection based on image data. One immediate application is to use SR to **enhance template matching for target detection**, i.e., given a grey/color image I and a target template X (for example, a face, a building or a military target, the template itself may be a simple binary image, a grey scale image, or a color image), the goal will be to search for the template in the entire image. To find the existence and position of X in the given noisy image Y, a template matching technique is often used; i.e., moving X on Y and attempting to find the **peak of a similarity metric such as the correlation function or mutual information**. This is similar to matched filtering in a single dimensional problem. As Y often appears to be a grey image, one would expect some error and considerable computational burden as we try to find the template X in the image Y directly when Y is a noisy image and/or the matching algorithms are sensitive to the small variations in the grey levels. The problem is actually similar to the image registration problem in which Syracuse University has considerable experience. Therefore, in this problem, quantization (or segmentation) of Y may become a very important preprocessing step. Furthermore, we would expect the SR effect (both in quantization and after quantization) to occur in a similar manner as in SR enhanced median filtering. The potential for improvement may be significant. A variety of applications can be pursued within this framework. These might include face recognition, image registration (to improve the performance via registration after segmentation) and may even be applicable to the single dimensional problem.

9.4 Stochastic Resonance in Image Visualization

a. Enhancement of RGB Imagery

In our prior work in Phase I, it has been shown that the quality of the thresholded 'Lenna' image can be improved by adding a suitable two-peak noise. In future studies, the optimal noise for a noisy image should be determined where the reference image is no longer available. To better utilize the spectral as well as the spatial information, some image models such as **multiple-dimensional Markov Random Field models** will be developed to better describe the correlations between R, G and B (red, green, blue) bands to further improve the image quality. Performance evaluation as well as adaptive learning algorithms will be pursued.

b. Multi- and/or Hyperspectral Imagery

In this area, the potential of improving Remote Sensing (RS) detection performance as well as classification accuracy via SR will be explored. Our preliminary results show that suitable Gaussian SR noise can improve the accuracy of the **Gaussian Maximum Likelihood Classifier** (GMLC). The role of SR in this area will be further examined utilizing the extensive techniques developed at Syracuse University [Varshney and Arora, 2004]. Possible enhancement of other processing algorithms by adding SR noise will also be evaluated. The form of suitable SR noise for specific algorithms will be derived.

c. Hyperspectral Image Visualization

Principal Component Analysis (PCA) and Segmented PCA (SPCA) based image visualization techniques [Tsagaris, 2005] are widely used to condense the information contained in hyperspectral images. However, due to the non-Gaussian nature of the hyperspectral images, PCA and SPCA may not fully utilize the hyperspectral imagery information. Moreover, PCA based techniques may introduce some artifacts. In general, a universal optimal method to display the contents of the hyperspectral images does not exist. Therefore, in future analyses, the potential enhancement by SR in various visualization schemes will be pursued and a better visualization result is expected.

d. Enhancement of Region of Interest (ROI)

Determination of the region of interest (ROI) is an essential part of many applications such as the determination of possible location of tumors in medical images and the possible target/incidents in video surveillance. For these types of problems, description of the statistical properties of the ROIs to specify the exact form of the target distribution is generally impossible, although some features may be available. In future studies, SR based techniques should be developed to better extract the features from the image and determine the location and size of the ROIs. This approach will have a very broad application in medical image processing.

9.5 Visual Image Fusion Considerations for Human Perception

The rapid development in imaging and computing technology has fostered utilization of visual information for situation assessment and decision making. Multiple image devices of either the same or different modalities are used to capture images. These images are then fused to integrate each image's information content to render a single image of enhanced quality. As image fusion techniques become available, evaluation of their performance is of high interest. However, there is a requirement for specific metrics to evaluate the quality of fused images and its affect on the human vision system (HVS). This problem has been addressed at Syracuse University [Chen and Varshney, 2006]. For the current proposal, such algorithms may be pertinent to the fusion of multiple images each containing an IID realization of added noise.

Comparative evaluation of fused images is a critical step to evaluate the relative performance of image fusion algorithms. Human visual inspection is often used to assess the quality of fused images. Here, we discuss some variants of a new image quality metric based on the human vision system (HVS). These measures evaluate the fused image quality by comparing its visual differences with the source image thus requiring no ground truth knowledge. First, the images are divided into several local regions. These regional images are then transformed to the frequency domain. Second, the difference between the transformed local regional images is weighted with a human contrast sensitivity function (CSF). The local regional image quality is obtained by computing the mean square error (MSE) of the weighted difference of the images obtained from the fused regional image and source regional image. ***Finally, the quality of a fused image is the weighted summation of the local regional image quality measures. Our experimental results show that these metrics are consistent with perceptually obtained results.***

The design of a universal objective measure for image quality is difficult due to the large variety of image fusion applications. Often, ground truth images are not available so that image quality evaluation is complicated. Since humans are the final users of fused images, human visual inspection is often used as the quality measure. However, human inspection may not always be possible due to large input and output data sizes and associated financial costs. Several automatic image evaluation algorithms have been developed recently. A performance measure for pixel-level fusion performance that compares the edge information of fused images with edge information of input images has been proposed [Xydeas, 2000]. Later, it was used to calculate the affect of noise on image fusion [Petrov, 2003]. Mutual information was also employed to evaluate fusion performance [Qu, 2002]. Wang and Shen [Wang, 2003] proposed a quantitative correlation analysis method to evaluate hyperspectral image fusion algorithms. [Piella, 2004] has proposed some new quality measures based on an image quality index proposed by Wang and Bovik [Wang, 2002]. ***However, there is no established direct relationship between these evaluation measures and the real perceptual results of humans.***

A number of linear and nonlinear models have been proposed to simulate the response of the HVS. Use of a contrast sensitivity function (CSF) to modify the difference of two images in the frequency domain is popular. However, as the CSF is only applied in the frequency domain, it does not include the nonlinear local spatial information such as local luminance and contrast. However, local spatial information is very important in image fusion because it strongly relates to the image content that is transferred from input images to the fused image. **Here, we propose**

a scheme to compare the quality of different fused images by comparing them with input images based on both types of local information. These are given by the salience of a set of localized windows and the difference in the frequency domain filtered by CSF. Compared to other measures, our proposed metrics have several advantages. First, we introduce this methodology to evaluate the image fusion performance. Second, in the conventional CSF based methods, the entire image is considered. We employ CSF based methods on a region-by-region basis. This is more suitable for the fusion application because one should examine image quality at a local level rather than at a global level. Moreover, image content and statistics vary over an image and, therefore, our region-based image quality measure is more appropriate. **Finally, compared to other methods such as mutual information based methods, our proposed methods require much less computation.**

We illustrate the utility of several image quality metrics under investigation at Syracuse University (SU) by applying them to evaluate image quality using three fusion schemes. The first is the ***'wavelet' fusion algorithm*** where the input images are decomposed using a length four Daubechies wavelet filter. The fused image coefficients are computed by choosing the corresponding input image coefficients with largest wavelet domain amplitude and by averaging the coefficients of lowest resolution. The number of decomposition levels is two. This fusion algorithm emphasizes the fused image edge information. The second algorithm is the ***'averaging' method*** consisting of the average of the input images on a pixel-by-pixel basis. The third is the ***'Laplacian' pyramid based method*** where the input image is decomposed using a Laplacian pyramid decomposition and the fused image is reconstructed by averaging the low resolution components and selecting the coefficients with largest amplitude for the remaining coefficients.

The images shown in Fig. 8.1 are created by blurring the original speckle noisy 'Lenna' image (a) of size 256x256 using a 10x10 mean filter twice to obtain two input images (b and c) to be fused. Input images (b) and (c) have blurring that occurs in the right and left half of the images, respectively. In Fig. 8.1d, e, and f, we observe that the 'wavelet' method outperforms the other two. Performance results for 10 repeated trials are shown in Table 1. ***These include new metrics developed at Syracuse University denoted by the first three superscripted Q's.*** For low SNR, the 'averaging' method is best, while at high SNR, the 'wavelet' method is superior. The results were confirmed by seven human subjects. They preferred the 'wavelet' method for SNR > 25dB and the 'averaging' method for SNR<25dB. Details of the new performance metrics are considered further in [Chen & Varshney 2006].

In Phase II, we will continue our consideration of new perceptually based quality metrics for image fusion which do not require a reference image. Compared to other measures, these metrics are easier to calculate and are also applicable to different input modalities. Experimental results show that our quality metrics fit the results of human visual inspection well and also correlate well with other measures. Further, our proposed metrics can be used to determine the fusion performance of different algorithms with different noise PDFs. Several extensions to this work such as extending our input images to include multi-color images, hyperspectral imaging, and stochastic resonance effects were initiated in the Phase I effort and will be continued in Phase II.

Fig. 9.1 Fused images for the speckle noisy 'Lena' image set with different SNRs. The left column uses the 'wavelet' method, the center column uses the 'averaging' method, and the right column uses the 'Laplacian' method. (a),(b), (c): fused results when SNR = 15dB; (d),(e),(f): SNR = 20dB; (g),(h),(i): SNR = 25dB.

SNR(db)	Fusion Method	Q^M	Q^D	Q^A	$Q_P^{X_1X_2/X_F}$	$M_{X_F}^{X_1X_2}$	RMSE	$I(X_R, X_F)$
15	Averaging	**208**	307	**137**	0.5703	**2.3658**	**16.7411**	**1.2364**
15	Wavelet	357	308	342	0.5730	1.7290	25.1462	0.9208
20	Averaging	**145**	269	**73**	.6137	**2.8120**	**12.4013**	**1.4701**
20	Wavelet	163	**218**	124	**.6493**	2.2712	15.1952	1.2615
25	Averaging	126	262	**51**	0.6409	**3.1972**	10.6843	**1.6805**
25	Wavelet	**103**	**196**	55	**0.7015**	2.7726	10.2462	1.6035
Human Inspection	7/7	5.5/7	5.5/7	6/7	5.5/7	3.5/7	6.5/7	4.5/7

Table 9.1 Performance results for visual image fusion and human visualization.

9.6 Stochastic Resonance in Distributed Systems

It has been shown in our preliminary work during Phase I that SR plays a role in decision fusion systems. In [Chen et al., 2006e], we have shown the existence of the SR effect in distributed detection systems for the two hypotheses detection problem. In future research, we will formulate and explore the general framework of SR in distributed detection and estimation problems. The role of SR in such a collaborative context forms a fascinating research area, one that has never been considered.

a. SR in Distributed Estimation

In distributed estimation systems, the data is compressed (often quantized) at the sensor before it is transmitted to the fusion center. As only limited information is transferred from the local sensors to the processing sensor, some nonlinear transformations are often applied to the received data. The role of SR noise in such applications will be evaluated and the possible estimation accuracy improvement by adding SR noise will be examined.

b. SR in Distributed Detection

In distributed detection problems, Likelihood Ratio Tests (LRT) or locally optimal detectors (LOD) (when signal strength is weak) are often used at the local sensors. These types of detectors are optimum when each sensor is conditionally independent [Varshney, 1996]. However, in many practical applications, sensors are at least partially dependent. In that case, the LRT or LOD are no longer optimum detection approaches. Although the optimal solution for this problem may be solvable when the full information about all the sensors is available, the design of each corresponding local detector without this knowledge (which is the desired goal of distributed systems) is often either too complicated or very difficult to implement. In this proposed study, we will explore the possible enhancement by adding SR noise to the local sensors utilizing the available (partial or full) information available about the local sensors to improve the local detection performance without altering the local detectors. Compared to the complicated approaches that are currently being investigated in the distributed detection literature, our SR based approach may provide a very simple and effective approach for improved detection performance. This will have a major impact on application areas such as vehicle health management (VHM) efforts for avionic systems currently being studied at Syracuse University in collaboration with NASA Langley Research Center.

REFERENCES

Ahumanda, A., "Simplified vision models for image quality assessment, *SID International Symposium Digest of Technical Papers*, 27 (1996) 397-400.

Albert, T.R., A.R. Bulsara, G. Schmera, M. Inchinosa, , "An evaluation of the stochastic resonance phenomenon as a potential tool for signal processing," *27th Asilomar Conf. on Signals, Systems and Computers*, Vol. 1, pp. 583-587, 1-3 Nov. 1993.

Ando, B. and S. Graziani, "Adding noise to improve measurement," *IEEE Instrumentation & Measurement Magazine*, pp. 24-31, March 2001.

Anishchenko, V.S., A.B. Neiman, F. Moss, L. Schimansky-Geier, "Stochastic resonance; noise-enhanced order," *Physics –Uspekhi*, **42** (1) 7 -36 (1999).

Asdi, A.S., and A. Tewfik, "Detection of weak signals using adaptive stochastic resonance," *1995 International Conference on Acoustics, Speech, and Signal Processing*, ICASSP-95, Vol. 2 , pp. 1332-1335, 9-12 May 1995.

Begas, J., "On the Statistical Analysis of Dirty Pictures," *J.R. Statist,* Soc. B 49, No. 3, Pp. 259-302, 1986.

Benzi, R., A. Sutera, A. Vulpiani, "The mechanism of stochastic resonance," *Journal of Physics A: Mathematical and General*, **14** (1981) L453 – L457.

Benzi, R., G. Parisi, A. Sutera, A. Vulpiani, "Stochastic resonance in climate change," *Tellus*, **(34)**, pp. 10-16, 1982.

Bingabr, M., and P.K. Varshney, "Recovery of corrupted DCT coded images based on reference information," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 14, No. 4, April 2004.

Bingabr, M., P.K. Varshney, and B. Farell, "An image quality measure for image communication," *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1, pp. 1019 – 1024, 5-8 October 2003.

Bingabr, M., and P.K. Varshney, "A novel error correction method without overhead for corrupted JPEG images," *IEEE Internatational Conf. on Image Processing*, Vol. 1, pp. I-245 – I-248, 22 -25 September 2002.

Bremaud, P., *Markov Chains, Gibbs Field, Monte Carlo Simulation and Queues,* Springer-Verlag, New York, 1999.

Castro, F.J., Kuperman, M.N., Fuentes, M., and Wio, H.S., "Experimental evidence of stochastic resonance without tuning due to non-Gaussian noise," *Physical Review E*, vol. 64, no. 051105, 2001.

Chair, Z., and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems,* AES-22: 98 - 101, 1986.

Chapeau-Blondeau, F., and D. Rousseau, "Noise-Enhanced Performance for an Optimal Bayesian Estimator," *IEEE Transactions on Signal Processing*, Vol. 52, No. 5, May 2004.

Chen, B., Ruixiang Jiang, T. Kasetkasem, and P.K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Transactions on Signal Processing*, 52(12): 3354 - 3358, 2004.

Chen, H., and P.K. Varshney, "A human perceptual inspired quality metric for image fusion based on regional information," *International Journal of Information Fusion*, accepted for publication, 2006.

Chen, H., P.K. Varshney, J.H. Michels, and S. Kay, "Approaching Near Optimal Detection Performance via Stochastic Resonance," *IEEE International Conference on Acoustics,*

Speech, and Signal Processing, Toulouse, France, May 14 – 19, 2006a, finalist in the Student Paper Contest.

Chen, H., P. Varshney, S. Kay, and J.H. Michels, "Theory of the Stochastic Resonance Effect in Signal Detection: Part I – Fixed Detectors," submitted to *IEEE Transactions on Signal Processing*, 2006b.

Chen, H., P. Varshney, J.H. Michels, and S. Kay, "Improving Nonparametric Detectors via Stochastic Resonance," *40th Annual Conference on Information Sciences and Systems*, Princeton University, 2006c.

Chen, H., P. Varshney, J.H. Michels, and S. Kay, "Performance Enhancement of Median Filters via Stochastic Resonance", to be submitted, 2006d.

Chen, H., P. Varshney, J.H. Michels, and S. Kay, "Can addition of noise improve distributed detection performance?," *9th International Conference on Information Fusion* (Fusion 2006) to be held in Florence, Italy, 10 – 13 July, 2006e.

Chen, H., *Ph.D. dissertation in progress*, Syracuse University, Syracuse, NY, 2006f.

Chiang, S-S., C-I. Chang , and I.W. Ginsberg, "Unsupervised target detection in hyperspectral images using projection pursuit", *IEEE Transactions on Geoscience and Remote Sensing*, **39**, no. 7, 2001, 1380-139

Collin, J.J., Chow, C.C., and Kiss, L., "Stochastic resonance without tuning," *Nature*, vol. 376, pp. 236 – 238, 1995.

Conte, E., Galati, G., Longo, M., "Exogenous Modeling of Non-Gaussian Clutter," *J. IERE*, **57**, (4): pp. 151-155, 1987.

Daly, S., The visible differences predictor: An algorithm for the assessment of image fidelity, in: A. B. Watson (Ed.), *Digital Images and Human Vision*, The MIT Press, Cambridge, MA, 1993, pp. 179-206.

Damera-Venkata, N., T. D. Kite, W. S. Geisler, B. L. Evans, A. C. Bovik, Image quality assessment based on a degradation model, *IEEE Transactions on Image Processing*, 9 (2000) 636-650.

Galdi, V. and V. Pierro, I.M. Pinto, "Evaluation of Stochastic-Resonance-Based Detectors of Weak Harmonic Signals in Additive White Gaussian Noise," *Physical Review* E, 57 pp. 6470 – 6479, 1998.

Gammaitoni, L., P. Hanggi, P. Jung, F. Marchesoni, "Stochastic resonance," *Review of Modern Physics*, Vol. 70, No. 1, pp. 223 – 287, January 1998.

Geman, S., and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesion Restoration of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. PAMI-6, No. 6, Pp. 721-41, November 1984.

Gilks, W., S. Richardson, D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, NY, 1996.

Godivier, X., and F. Chapeau-Blondeau, "Stochastic resonance in the information capacity of a nonlinear dynamic system," *International Journal of Bifurcation and Chaos*, vol. 8, no. 3, pp. 581 – 589, 1998.

Goychuk, I. and P. Hanggi, "Stochastic resonance in ion channels characterized by information theory," *Physical Review E*, vol. 61, no. 4, pp. 4272 4280, 2000.

Hasting, W.K., "Monte Carlo Sampling Method Using Markov Chains and Their Applications," *Biometrika,* Vol. 57, No. 1, Pp. 97-109, Apr. 1970.

Inchiosa, M.E., and A.R. Bulsara, "Signal detection statistics of stochastic resonators," *Physical Review* E, 53 R2021 – R2024, 1996.

Jia, X., and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote sensing image display and classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 37, pp. 538-542, Jan. 1999.

Jung, P., and K. Wiesenfeld, "Too quiet to hear a whisper," *Nature*, Vol. 385, p. 291, 23 January 1997.

Kasetkasem, T., and P.K. Varshney, "Statistical Characterization of cluster scenes based on a Markov random field model," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 39, No 3, pp. 1035 - 1050, July 2003.

Kasetkasem, T., and P.K. Varshney, "Clutter Patch Identification Based on a Markov Random Field Model", *Proceedings of IEEE Radar Conference 2002*, Long Beach, CA, April 2002.

Kassam, S.A. and J.B.Thomas, "Dead-zone limiter: An application of conditional tests in nonparametric detection," *The Journal of the Acoustical Society of America*, vol. 60, pp. 857 – 862, 1976.

Kassam, S.A. and J.B.Thomas (Eds.), *Nonpaprametric Detection: Theory and Applications*, Dowden, Hutchinson and Ross, 1980.

Kay, S., *Fundamentals of Statistical Signal Processing, Vol. II*, Prentice-Hall, Englewood Cliffs, NJ, 1998.

Kay, S., "Can detectability be improved by adding noise?," *IEEE Sig. Proc. Letters,* (7), No. 1, pp. 8-10, Jan. 2000.

Kay, S., A. Nuttall, P. Baggenstoss, "Multidimensional Probability Density Function Approximations for Detection, Classification, and Model Order Selection," *IEEE Trans on Signal Proc.*, Vol 49, pp. 2240-2252, Oct. 2001.

Kay, S., "Alternative Derivation for an Optimal Stochastic Resonance Detection Framework," Memorandum 2005-10-19 for JHM Technologies, 19 October 2005.

Kay, S., "Optimal Decision processing by the Transformation Method," Memorandum 2006-3-2 for JHM Technologies, 2 March 2006a.

Kay, S., J.H. Michels, H. Chao, P. Varshney, "Reducing Probability of Error using Stochastic Resonance," to be published, *IEEE Signal Processing Letters*, 2006b.

Kosko, B., and S. Mitaim, "Stochastic Resonance in Noisy Threshold Neurons," *Neural Networks*, **16**, 755-761, 2003.

Kosko, B., and S. Mitaim, "Robust stochastic resonance for simple threshold neurons", *Physical Review E*, vol. 70, no. 031911, 2004.

Li, H., and J.H. Michels, "Parametric Adaptive Modeling for Hyperspectral Imaging", *International Conference on Acoustics, Speech, and Signal Processing* (ICASSP), Montreal, CA, 17 – 21 May 2004.

Lindner, J.F., B.K. Meadows, W.L. Ditto, M.E. Inchiosa, A.R. Bulsara, "Array enhanced stochastic resonance and spatiotemporal synchronization," *Physical Review Letters*, Vol. 75, No. 1, pp. 3 – 6, 3 July 1995.

Lipshitz, S.P., R.A. Wannamaker, and J. Vanderkooy, "Quantization and dither: A theoretic survey," *J. Audio Eng. Soc.*, Vol. 40, no. 5, pp. 355-374, May 1992.

Liu, B., B. Chen, J. Michels, "A GLRT for radar detection in the presence of compound-Gaussian clutter and additive white noise," *2nd IEEE Sensor Array and Multichannel Signal Proc. Workshop*, pp. 87-91, Aug. 2002.

Mannos, J., D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Transactions on Information Theory*, 20 (1974) 525-536.

Marks, R.J., B. Thompson, M.A. El-Sharkawi, W.L.J. Fox, R.T. Miyamoto, "Stochastic Resonance of a threshold detector: image visualization and explanation," IEEE International Symposium on Circuits and Systems, Vol. 4, pp. IV - 521 – IV - 523, 26 – 29 May 2002.

Marroquin, J., S.Mittter, and T. Poggio, "Probabilistic Solution of Ill-Posed Problem in Computer Vision," *Journal of the American Statistician Association,* Vol. 82, No. 397, Pp. 76-89. Mar. 1987.

McDonnell, M.D., N.G. Stocks, C.E.M. Pearce, and D. Abbott, "Stochastic resonance and data processing inequality," *Electronics Letters*, Vol. 39, No. 17, pp. 1287-1288, August 2003.

Michels, J.H., M. Rangaswamy, and B. Himed, "Performance of parametric and covariance based STAP tests in compound-Gaussian clutter," *Digital Signal Processing*, Vol. 12, No. 2/3, April/July 2002.

Michels, J.H., B. Himed, M. Rangaswamy, "Performance of STAP tests in Gaussian and Compound-Gaussian Clutter," *Digital Signal Processing*, Vol. 10, No. 4, October 2000a.

[Michels, J.H., M. Rangaswamy, B. Himed, "Robust STAP Detection in Discrete Airborne Radar Clutter", *2001 Defence Applications of Signal Processing (DASP) Workshop*, Barossa Valley, Australia, 16-21 Sept., 2001a.

Michels, J.H., B. Himed, M. Rangaswamy, "Further Evaluations of STAP Tests in Compound-Gaussian Radar Clutter," *Proc. 9th Annual Adaptive Sensor Array Processing Workshop*, Lexington, MA, 13-14 March 2001b.

Michels, J.H., M. Rangaswamy, B. Himed, "Performance of STAP Tests in Compound-Gaussian Clutter," *IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, 16-17 March 2000b.

Mitaim, S., and B. Kosko, "Adaptive stochastic resonance," Proc. IEEE, Vol. 86, No. 11, pp. 2152-2183, Nov. 1998.

Mitaim, S., and B. Kosko, "Adaptive stochastic resonance in noisy neurons based on mutual information," *IEEE Transactions on Neural Networks*, vol. 15, no. 6, pp. 1526 – 1540, November 2004.

Muller, H.-P., E. Kraft, A. Ludolph and S.N. Erne, "New methods in fMRI analysis," *IEEE Engineering in Medicine and Biology*, pp. 134-142, Sept/Oct. 2002.

Nicolis, C., G. Nicolis, "Stochastic aspects of climate transition – additive fluctuations," *Tellus*, Vol. 33, pp. 225-237, 1981.

Niu, R., B. Chen, and P.K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channel in wireless sensor networks," *IEEE Transactions on Signal Processing*, March 2006.

Qu, G., D. Zhang, P. Yan, "Information measure for performance of image fusion," *Elect. Ltrs,* 38 (2002) 313-315.

Ozturk, A., and E.J. Dudewicz, "A New Statistical Goodness-of-Fit Test Based on Graphical Representation," *Biom. J.,* Vol. 34, No. 4, Pp. 403-427, 1992.

Papadopoulos, H.C., and G.W. Wornell, "A class of stochastic resonance systems for signal processing applications," 1996 IEEE ICASSP-96, Vol. 3, pp. 1617 – 1620, 7-10 May 1996.

Petrovic, V.S., C. S. Xydeas, "Sensor noise effects on signal-level image fusion performance," *International Journal of Information Fusion*, 4 (2000) 167-183.

Picinbono, B., "On deflection as a performance criterion in detection," *IEEE Transactions on Aerospace and Electronic Systems*, 31(3): 1072 – 1081, July 1995.

Piella, G., "New quality measures for image fusion," *Proceedings of the 7th International Conference on Information Fusion*, Stockholm, Sweden, 2004, pp. 542 - 546.

Plesser, H.E., and S. Tanaka, "Stochastic resonance in a model neuron with reset," *Physics Letter A*, 225 (1997) 228 – 234, 3 February 1997.

Rangaswamy, M., et. al., "Signal Detection in Correlated Gaussian and Non-Gaussian Radar Clutter," Rome Laboratory (now AFRL) Final Technical Report, RL-TR-93-79, May 1993.

Rangaswamy, M., J.H. Michels, B. Himed, "Statistical analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds," *Proc. of the IEEE Radar Conf.*, Long Beach, CA, pp. 304 – 310, 22-25 April 2002.

Repperger, D.W., et. al. "Application of data mining and knowledge discovery techniques to enhance binary target detection and decision-making for compromised visual images," to be published.

Repperger, D.W., et. al. "Optimization of Stochastic resonance with Application to Object/Image Enhancement," *submitted for publication.*

Robila, S.A., and P. K. Varshney, "Target detection in hyperspectral images based on independent component analysis," *SPIE AeroSense*, Orlando, FL, April 2002.

Robila, S.A., P.K. Varshney, "Further Results in the use of Independent Component Analysis for Target Detection in Hyperspectral Images," *Proc. SPIE Automatic Target Recog.*, XIII, vol. 5094, pg. 186-195, Apr. 2003.

Ruanaidh, O., *Numerical Bayesian Methods Applied to Signal Processing*, Springer Verlag, NewYork, 1996.

Rushmeier, H., G. Ward, C. Piatko, P. Sanders, B. Bust, Comparing real and synthetic images: Some ideas about metrics, *6th Euro graphics Workshop on Rendering*, 1995, pp. 213-222.

Saha, A., G.V. Anand, "Design of detectors based on stochastic resonance," *Signal Proc.*, 83, 1193 – 1212, 2003.

Shah, C.A., M. K. Arora, and P. K. Varshney, Unsupervised classification of hyperspectral data: an ICA mixture model based approach, *International Journal of Remote Sensing*, vol. 25(2), pp. 481 - 487, Jan. 2004.

Shah, C.A., and P. K. Varshney, "A higher order statistical approach to spectral unmixing of remote sensing imagery," *IEEE International Geoscience and Remote Sensing Symposium, Exploring and Managing a Changing Planet*, Anchorage, AK, September 20-24, 2004.

Shah, C.A., P. Watanachaturaporn, M. K. Arora, and P. K Varshney, "Some Recent Results on Hyperspectral Image Classification", *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, NASA Goddard Spaceflight Center, Greenbelt, MD, October 27-28, 2003.

Shah, C.A., M. K. Arora, and P. K. Varshney, "ICA mixture model for unsupervised classification of non-Gaussian classes in multi/hyperspectral imagery," *Proc. SPIE AeroSense, Intern. Soc. for Opt. Eng*, Orlando, FL, 21-25 April, 2003.

Sheikh, H.R., A. C. Bovik, and L.K. Cormack, "No-Reference Quality Assessment using Natural Scene Statistics: JPEG2000," *IEEE Transactions on Image Processing*, Vol. 14, Issue 11, pp. 1918 – 1927, November 2005

Simonotto, E., M. Riani, C. Seife, M. Roberts, J. Twitty, F. Moss, "Visual perception of stochastic resonance," *Physical Review Letters*, Vol. 78, No. 6, pp. 1186-1189, 10 February. 1997.

Slamani, M.A., "A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume," *Ph.D. Dissertation,* Syracuse University, Syracuse, NY, May 1994.

Stocks, N.G., "Suprathreshold Stochastic resonance in multilevel threshold systems," *Physical Review Letters*, Vol. 84, No. 11, pp. 2310-2313,13 March 2000.

Sutton, B.P., D.C. Noll and J.A. Fessler, "Fast, iterative, image reconstruction for MRI in the presence of field inhomogeneities," *IEEE Trans. on Medical Imaging*, Vol. 22, No. 2, pp. 178-188, February 2000.

Tanner, M., *Tools fo r Statistical Inference*, Springer Verlag, New York, 1993.

Thomas, J., "Nonparametric detection," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 623 – 631, May 1970.

Tsagaris, V., V.Anastassopoulos and G.A. Lampropoulos, "Fusion of hyperspectral data using segmented PCT for color representation and classification", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, pp. 2365-2375, October 2005.

Tougaard, J., "Signal Detection Theory, Detectability and Stochastic Resonance Effects," *Biolog. Cybernetics*, **87**, 79–90, 2002.

Tyo, A.J., A. Konsolakis, D. Diersen, and R. C. Olsen, "Principal-components based display strategy for spectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, pp. 708-718, March 2003.

Vapnik, Vladimir N., "Statistical Learning Theory", Wiley-Interscience, 1996.

Varshney, P.K., and M.K. Arora (Eds.), *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*, Springer Verlag, 2004

Vaudelle, F., J. Gazengel, G. Rivoire, X. Godivier and F. Chapeau-Blondeau, "Stochastic resonance and noise-enhanced transmission of spatial signals in optics: the case of scattering," *J. Opt. Soc Amer. B*, Vol. 15, No. 11, pp. 2674 – 2680, November 1998.

Wang, O., Y. Shen, Y. Zhang, J. Q. Zhang, "A quantitative method for evaluating the performances of hyperspectral image fusion," *IEEE Transaction on Instrumentation and Measurement*, 52 (2003) 1041-1047.

Wang, Z., A.C. Bovik, and B.L. Evans, "Blind measurement of blocking artifacts in images," *Proceedings of the International Conference on Image Processing*, Vol. 3, pp. 981 -984, Vancouver, Canada, September 10 – 13, 2000.

Wang, Z., A. C. Bovik, A universal image quality index, *IEEE Signal Processing Letters*, 9 (2002) 81-84.

Wang, Z., A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, 13, pp. 600-612, (2004).

Wannamaker, R.A., S.P. Lipshitz, J. Vanderkooy and J.N.Wright, "A theory of nonsubtractive dither", IEEE trans. On Signal Processing, volume 48, pages 499-516. Feb. 2000.

Wiesenfeld, K., and F. Jaramillo, "Minireview of stochastic resonance," *Chaos*, (8), No. 3, pp. 539–548, Sept. 1998.

Winkler, G., *Image Analysis, Random Field and Dynamic Monte Carlo Methods,* Springle-Verlay, New York, 1995.

Xu, B., J. Li, and J. Zheng, "How to tune the system parameters to realize stochastic resonance," *Journal of Physics A: Mathematical and General*, **36** (2003) 11969 – 11980.

Xydeas, C., V. Petrovic, "Objective image fusion performance measure," *Electronics Letters* 36 (2000) 308-309.

Yao, K., "A representation theorem and its applications to spherically invariant random processes," IEEE transactions on Information theory, IT-19, pp. 600-608, 1973.

Ye, Q., H. Huang, X. He, C. Zhang, "A SR-based radon transform to extract weak lines from noise images," *Proc. of the 2003 International Conference on Image Processing*, Vol. 1, pp. I-849-852, 14-17 Sept. 2003.

Zhang, Z., R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, Vol. 87 of, 1999, pp. 1315-1328.

Zhou, H., M. Chen, M. F. Webster, Comparative evaluation of visualization and experimental results using image comparison metrics, in: Visualization, *Proceedings of IEEE*, 2002, pp. 315-322.

Zozor, S., and P.O. Amblard, "Stochastic resonance in discrete time nonlinear AR(1) models," *IEEE Trans. on Signal Proc.*, Vol. 47, No. 1, January 1999.

Zozor, S., and P.O. Amblard, "On the use of stochastic resonance in sine detection," *Signal Processing*, vol. 7, pp. 353 – 367, March 2002.

Zozor, S., and P.O. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Transactions on Signal Processing*, Vol. 51, No. 12, pp. 3177-3181, December 2003.